IFKAD
International Forum on Knowledge Asset Dynamics

21st International Forum on Knowledge Asset Dynamics

Book of Papers

presented at

IFKAD 2026

Intelligent Knowledge for Sustainable Organizations

1-3 July 2026
Corvinus University of Budapest
Budapest, Hungary

21st International Forum on Knowledge Asset Dynamics

BOOK OF PAPERS

presented at

IFKAD 2026

Intelligent Knowledge for Sustainable Organizations

1-3 July 2026
Corvinus University of Budapest
Budapest, Hungary

ISBN 978-88-96687-20-8

This Book of Papers contains all presented papers, non-reviewed, and has been distributed during IFKAD 2026 conference, held at the Corvinus University, in Budapest (Hungary) from 1 to 3 July 2026

INDEX

10 *Maximilian Broglio*
AI as a Driving Force for Sustainability, Resilience, Regeneration and Inclusiveness in the Context of Smart Cities

18 *Giovanna Ferraro, Ivana Quinto, Giuseppe Scandurra, Antonio Thomas*
Artificial Intelligence and Sustainability Management as Innovation Strategies for Sustainable Competitiveness: Evidence from Innovative SMEs

26 *Maria Cristina Manocchio, Francesca Di Virgilio, Gianluigi Mangia, Antonia Puccio*
AI-Enabled Knowledge Ecosystems: Conditions for Learning and Deskilling Across Economic Sectors

36 *Juan-Gabriel Cegarra-Navarro, Aurora Martínez-Martínez, Scott Erickson, Jorge Cegarra-Sánchez*
When Knowledge Hiding Enhances Self-Assessment in Multiple Sclerosis

43 *Matteo Meacci, Antonella Martini, Gabriele Goretti, Caterina Dastoli*
Beyond Single Tasks: How Students Interact with AI in Design Processes

52 *Suncana Tuksar, Vanja Bevanda*
Building Trust in Hybrid Teams: The Role of AI-Assisted Storytelling

61 *László Berényi*
ChatGPT Acceptance among Higher Education Students: A Simplified Model

69 *Tamás Kristóf*
Advancing Foresight and Futures Literacy by Generative Artificial Intelligence

78 *Maayan Nakash*
Beyond Solutionism: A Pre-Implementation Diagnostic Framework for Strategic Knowledge Management Alignment

86 *Hairong Zhai, Farag Edghiem*
Emotional Knowledge and Manipulative Design in AI-Powered Fashion Recommendation Systems

94 *Anikó Csepregi, Coşkun Akça*
Project Tailoring as a Knowledge Legitimacy Process in Hybrid Intelligence Environments

103 *Filomena Pagnozzi, Helen Zhao*
Rethinking Help-Seeking in AI-Mediated Workplaces: Organizational and Individual Antecedents Through a Socio-Technical and Resource-Based Perspective

111 *Behrooz Moradi, Ettore Bolisani, Enrico Scarso, Kathrin Kirchner, Furong Cai*
AI–Gamification Role Matrices in Knowledge Management: Evidence and Research Gaps

121 *Micol Di Vita, Andrea Panelli*
Decision-Making Approaches and Entrepreneurs' Pivot Decisions? An RCT Experiment

127 *Francesca Loia, Irene Fulco, Barbara Aquilani, Marcello Martinez*
Artificial Intelligence Ambidexterity and Sustainable Business Models: A Configurational Case Study

136 *Andrea Gargiulo, Daniele Leone, Salvatore Scognamiglio, Petrella Raffaele Clemente*
Hybrid-Knowledge Management for Sustainable Competitiveness: AI-Human Synergy as Value Enabler

142 *Guan-Ting Chen, Sirirat Sae Lim*
The Key Factors in RTO Spin-off Development: Reconfiguring Organizational Knowledge for Spin-off Ventures

150 *Wan-Jung Chung, Jia-Bo Guo, Hao-Chieh Lin*
The Impact of Networking Skills on Quiet Quitting and Knowledge Hiding: Mediating Role of Job Crafting and Moderating Effect of Mindful Leadership

158 *Nicola Capolupo, Miriana Ferrara*
Knowledge Dissemination Strategies for Circular Economy: Evidence from Italian Local Governments

167 *Irene Spada, Antonella Martini, Gualtiero Fantoni*
Sustainability and Project Management: Discovering Skills Relations to Support Knowledge Intensive Organizations

177 *Demetrio Alessandro Trunfio, Ferdinando Potenza, Anna Mezzapesa, Carl Harvey Miguel, Giovanna Rotella*
Mentoring as a Sociodigital Ecosystem: Knowledge Sharing, Digital Communities and Student Support in an Engineering Department

INDEX

187 *Coşkun Akça, Albena Antonova, Anikó Csepregi, Zsuzsa Darida*
The Moderating Role of AI Self-Efficacy on the Effect of AI Anxiety on Unemployment Anxiety Among University Students: A Cross-Country Study

196 *Giulia Lippi, Luca Colomba, Paolo Garza, Enrica Vesce*
MatchIS: A Digital Platform Designed to Address Critical Gaps in Industrial Symbiosis Marketplaces

205 *Mervi Hokkanen*
The Roles of Art in Knowledge Creation in Urban Planning

214 *Clara Filosa, Lara Agostini, Anna Nosella*
Using Intelligent Knowledge to Offer Digital Services: a Lean Startup Interpretation

222 *Vittoria Magrelli, Isotta Vazzoler, Dario Natale Palmucci, Luca Giraldi*
Leading Intelligent Knowledge Management for Data-Driven Sustainable Innovation: Micro-Practices that Build Human–AI Complementarities

227 *Marco Scatto, Chiara Grosso, Karen Venturini, Alberto Petroni*
Mapping Green and Digital Strategy Adoption in the Packaging Industry: An Empirical Knowledge-Based Analysis

233 *Sergio Conterjnic, Cristina Simone, Antonio Laudando*
When traditional knowledge meets digitalization: the Italian High-Quality Olive Oil Value Chain and the platformized markets.

239 *Domenico Camarda, Maria Rosaria Stufano Melone*
Knowledge Management in Sustainable Energy Transitions: An Exploratory Agent-Based Approach

248 *Mercedes Grijalvo, Joaquín Ordieres-Meré, Lorna Anguilano*
Managing Circular Innovation Ecosystems through Knowledge Recombination: Assessing Ecosystem Health and ESG Outcomes in Phytoremediation-Based Circular Economy Models

258 *Nikolett Deutsch, László Berényi*
Economic Potential of ESCO Financing Models in Public Lighting: Evidence from Hungary

267 *Vincenzo Maione, Cristina Ponsiglione, Simonetta Primario*
Path Development and the Twin Transition in Europe

276 *Claudio D'Alonzo*
Knowledge Management and Artificial Intelligence: Governance and Liability in Business Management

283 *Dora Horvath*
Between Heritage and Change: Value-Oriented Digital Transformation in European Theatres

290 *Anna Maria Crespino, Marco Lucio Sarcinella, Lorenzo Quarta*
UX-Driven Digital Architecture for Knowledge Integration in Social-Ecological Systems

300 *Mariangela Lazoi, Marianna Lezzi, Nicolò Gianmauro Totaro*
A Multidimensional KPI-Based Framework for Evaluating AI-Driven Cybersecurity Solutions

308 *Stefano Marciano, Raffaele Fiorentino, Rosita Capurro, Francesca Peluso*
Territorial Development as Resource for Firm Growth Strategies: a Conceptual Framework

316 *Miklós Verseghi-Nagy, András Nemeslaki*
Measuring the Unmeasurable? A Systematic Literature Review of University Technology Transfer Efficiency

325 *Giovanni Felle, Antonio Arcadi, Davide Sergio Storelli, Valentina Chetta, Rosaria Lovecchio, Elisa De Bellis*
A Hybrid RAG Architecture for Knowledge Management in Urban Digital Twins

334 *Stefano Abbate, Anna Scotto di Santolo*
A Knowledge-Based Life-Cycle Framework for Sustainable Restoration Decision-Making

341 *Antonio Sonetto, Silvia Tommaso, Antonio Ricciardi*
Opening the Black Box of Incubation: Governance Archetypes and the Moderating Role of Regional Innovation Ecosystems in Start-up Growth

350 *Antonio Sonetto, Silvia Tommaso, Patrizia Pastore, Antonio Ricciardi*
Business Incubation as a Support Mechanism for the Circular Transition: Evidence from Italy

INDEX

358 *Luca Simone Macca, Alain Daou*
Learning to Scale: What Mentoring Really Does for Startups

364 *Gaetano Rago*
Accounting-Mediated Knowledge Management in Agrifood Supply Chains: Insights from Participant Observation on How Quality, Traceability and Cost Knowledge is Produced and Coordinated

371 *Anett Dobos, Zoltán Soma Kárpáti, Márk Recskó*
How Does Innovation Activity Impact Firm Competitiveness? Evidence from Hungarian Firms

377 *Andrea Comandé, Olga Ferraro*
Accounting Standardization and Circular Knowledge Frictions in Cooperative Banks

384 *Navini Perera, Erika Jáki*
Intelligent Knowledge Management Systems for Human-AI Collaboration in Financial Decision-Making: A Behavioural Finance Perspective on Value Creation Beyond Profit

393 *Ingrida Girnienė, Laura Valiukevičienė*
Assessing Knowledge Management Maturity in Public Sector Organizations: Model Development and Case-Based Validation

400 *Silvia Martelo-Landroguez, Jerónimo García-Fernández, Gabriel Cepeda-Carrión*
The Role of Servant Leadership and Technology in Driving Organizational Citizenship Behaviour

409 *Ana Lidia Quintero Ramírez, Gabriela Citlalli López-Torres, Silvia Alexandra Peña Gonzalez, Hannia Guadalupe García Vega, Francisco Javier Álvarez-Torres*
Sociodigital and Cultural Skills as Mediators of Technology Adoption in Higher Education Learning Ecosystems

417 *Silvana Revellino*
Beyond Profit, Through Data: How AI-Driven Academic Spin-offs Enact Public Value in Knowledge Ecosystems

425 *Sharmila Gope, Claudio Pascarelli, Mattia Mangia, Giovanni Coppini, Ivan Federico, Salvatore Causio, Marco Boetti, Federica Stella Blasi*
Towards an Integrated Digital Twin Maturity Assessment Framework for Coastal Cultural Heritage Resilience

434 *Carlo Drago, Angelo Leogrande, Valentina Giglio Moro, Linda Stella Guajardo Rios*
Artificial Intelligence and the Transformation of Knowledge Management Systems: A Consensus Community-Detection Bibliometric Analysis

440 *Attila Wieszt, Noémi Szilvia Lőrinczz, Amitabh Anand*
Governing Adaptation: Family Councils, Dynamic Capabilities, and Company Performance in Family Firms

448 *Viktoriia Apalkova*
Analytics Maturity and AI Governance for Organizational Trust: A Swiss Capability Perspective

455 *Gabriele Rizzo, Daniele Allega*
Redefining Financial Advisory: Human-Machine Collaboration and Organizational Transformation Under the EU Financial Data Access Regulation

464 *Leani Lauermann Koch, Aline Luiza Brusco Pletsch, Graciele Tonial, Florinda Matos, Beatriz Carla Koch*
Public Policies and Knowledge Democracy in Innovation Ecosystems: A Quintuple Helix Analytical Framework

472 *Anikó Almási*
The Organizational Challenges of Digital Transformation at SMEs

481 *Judit Fortvingler*
Microfoundations of Dynamic Capabilities: Intelligent Knowledge for Sustainable Internationalisation in Higher Education

490 *Ines Ben Atitallah, Remsei Sándor, Heba F. Zaher*
Human–LLM Collaboration in Tactical Production Decision-Making: A Literature Review

499 *Eduardo Trauer*
Co-Creation of Knowledge in Marketing: Human-AI Collaboration in the Construction of Proto-Personas

508 *Sara El Matouk, Daniele Allega*
Beyond Prediction: Building Sustainable Human–AI Knowledge Ecosystems in Healthcare Organizations

INDEX

515 *Daniele Allega, Gabriele Rizzo*
From Paradoxes and Entropies to Sustainable Knowledge Ecosystems: Architectural and Governance Principles

522 *Riccardo Michele Colangelo, Eleonora Petrazzuoli, Maria Menshikova*
Organisational AI Literacy and Readiness: Enabling Factors and Barriers

532 *Giovanna Afeltra, Maria Menshikova, Maria Pia Ciano*
New frontiers in Organizational Aesthetics

540 *Stefano Allega, Daniele Allega, Sara El Matouk, Giovanni Spatola*
AI-Mediated Knowledge Curation in Platform-Based E-Commerce

547 *Mauro Romanelli*
Smart Healthy Cities Advancing

552 *Alexandra Zbuchea, Monica Bira, Mauro Romanelli*
Cultural Organizations and Young Audiences: A Collaborative View through Artificial Intelligence

558 *Paola Demartini, Alessandro Salvatori*
The Digital Platform as Cognitive Infrastructure: Investigating Human-AI Synergy and Algorithmic Governance in Investment Management

567 *Maria Rosaria Marcone*
Digital Innovations in Sustainable Supply Chain Business Models

575 *Noémi Szilvia Lőrincz, Márta Kiss*
The Cascade Paradox: Information Refraction in Hierarchical Organisations, AI-Driven Power Shifts, and the Future of Knowledge Cascades

585 *Eduardo Sánchez-García, Javier Martínez-Falcó, Bartolomé Marco-Lajara, Matilda Shini, Grazia Dicuonzo, Zhadyra Konurbayeva, Assel Sadenova*
Knowledge-Based Mechanisms of Sustainable Competitive Advantage under Environmental Pressure

593 *Giulio Paolo Agnusdei, Donato Longo, Zahra Ardakani, Pier Paolo Miglietta*
Integrating Farmers' Behaviors into Agricultural Index-Based Insurance: A Knowledge-Based Theoretical Framework

602 *Fabrizio Amarilli*
The Impact of Knowledge Asymmetries on Business–IT Alignment Dynamics

611 *Nina Gerges, Paulo Selig, Gregorio Varvakis*
A KM-Based Approach to Service Value Co-Creation and Co-Destruction

618 *Olufemi Afolabi, Oluseye Jegede*
Artificial Intelligence and Marketing Creativity: A Thematic Synthesis of Disruption, Human–AI Collaboration, and Strategic Transformation

629 *Mariavittoria Marinucci, Giovanni Schiuma, Roberto Sala, Fabiana Pirola, Fabio Floreani*
Making Knowledge Visible: The Role of Process Mapping in Enabling Knowledge Management in Industrial SMEs

638 *Mattia Barsanti, Salvatore Tallarico, Luisa Pellegrini, Riccardo Dulmin, Simone Lazzini, Marco Praticò*
AI-Supported Decision-Making in Medical Device Procurement: From Ad Hoc Criteria to Meta-Criteria and Value Configurations

647 *Teresa Anna Rita Gentile, Ernesto De Nito*
Training Project Managers through Games and Simulations: A Literature Review

656 *Daria Podmetina, Anne Berthinier Poncet, Peter Robbins, Jelena Hartsenko, Jani Kiljala*
Knowledge as Experience: Embodied Artistic Practices, AI, and the Governance of Intelligent Knowledge for Sustainable Innovation

665 *Fabiola Colmenero Fonseca, Daniel Ruíz López, Ramiro Rodríguez Pérez, Lauren Yolanda Gómez Zamorano, Amparo Borrell*
Smart Knowledge for Sustainable Artisanal Production in the Integration of Science, Technology and Local Knowledge in Rural Communities in Guanajuato

INDEX

674 *Juan Francisco, Palomino Bernal, Fabiola, Colmenero Fonseca, Rosa María, López Salinas, Elsa Ivette, Aldana Ornelas, Javier, Cárcel Carrasco*
Intelligent Knowledge Ecosystems for the Dual Transition, Curricular Innovation, and Sustainable Production in Artisan Communities of Mexico

684 *Constantin Bratianu, Ruxandra Bejinaru*
The New Organizational Knowledge Dynamics Model that Integrates Human and Artificial Knowledge

690 *Nima Taraghi, Kathrin Kirchner, Ettore Bolisani, Tomas Cherkos Kassaneh, Enrico Scarso*
Generative AI in Fashion: How Marketing Experts Integrate GenAI in their Knowledge Intensive Work

699 *Elina Pentjusa, Elina Gaile-Sarkane*
The Hidden Cost of Flexibility: Knowledge Visibility in Flexible Work Arrangements

707 *Maria Vittoria Prudentino, Paola Canestrini, Federica Bandini*
The Possible Effects of Digital Transformation on Social Enterprises and NGO

714 *Francesca Dal Mas, Ludovica Schiavone, Giuseppe Pezzotti, Alberto Cavazza*
Integrating Knowledge and Relational Capital to Improve Elderly Care: An Interventionist Approach

719 *Emilia Romeo, Valerio Giampaola, Giuliano Marolla, Alberto Ferraris, Angelo Rosa*
Beyond Facilitation: Ethical and Behavioral Tensions in Technology-Enabled Knowledge Management

726 *Fabrizio Flavio Baldassarre, Claudio Rosito, Valerio Buono, Gennaro Durante, Lucrezia Maria De Cosmo*
An Index-Based Approach to SME Sustainability Learning and Governance

734 *Lorella Cannavacciuolo, Roberta De Cristofaro, Vincenzo Maione, Cristina Ponsiglione, Simonetta Primario*
An Integrated Framework for Evaluating Telemedicine Systems in Regional Healthcare Contexts

741 *Manuela Paolini, Domenico Raucci, Giuseppe Modarelll*
Managing Knowledge for Including Persons with Disabilities in Organizational Contexts

748 *Silvia Oliveira, Nathalia Werlang*
Absorptive Capacity in Complex Public Organizations: The Role of AI

756 *Shuyun Zhao, Krishna Venkitachalam*
Influence of Artificial Intelligence (AI) on Organizational Knowledge Creation (OKC) in Workplace Meetings

764 *Davide D'Emiliano, Laura Nucera, Andrea De Mauro*
Beyond Marketplace Reviews: Transformer-Based Sentiment Analysis of YouTube Influencer Ecosystems

772 *Harald Grundner*
Metrics and Value of Intelligent Knowledge: How Organizations Turn AI Options into Decision Capability, and Become More Sustainable

779 *Annabella Lattarulo*
Artificial Intelligence and Labour Law: The Employer's Obligation to Reskill Workers in the Digital Era

789 *Antonio Cimino, Vincenzo Corvello, Luca Famà, Ciro Troise*
Artificial Intelligence as Driver of Sustainable Performance in Entrepreneurial Ecosystems

797 *Maria Teresa Bianchi, Raffaele De Socio, Federica Mastrantonio*
Innovation and Sustainability as Judged by the Market

804 *Fabio Liberti*
FedHR5.0: A Human-Centric Federated Learning Framework for Organisational Learning and Workforce Competency Development in Healthcare

813 *Antonio Bassi*
INSPIRE PM as a Human–AI Hybrid Knowledge Ecosystem for Sustainable Project-Based Organizations

820 *Furong Cai, Ettore Bolisani, Kathrin Kirchner, Tomas Cherkos Kassaneh, Enrico Scarso, Behrooz Moradi*
AI-Driven Green Knowledge Management Processes and Startup Sustainable Performance: A Multidimensional Conceptual Framework

828 *Serena Strazzullo, Sara Ianniello, Jiyeonb Lee, Ilkyeong Moon*
Digital ESG Communication in Sustainable Fashion Consumption: Comparing Italy and South Korea

837 *Danilo Martino, Giulia Caso, Guglielmo Elefante, Claudio Pascarelli, Gabriele Papadia, Mariangela Lazoi*
TRACING: A Design-Oriented Framework for Enabling Governance Mechanisms in Cold Chain Networks

INDEX

846 *Faisal Rasool, Gustavo Morales-Alonso, Vicky O'Rourke, Antonio Hidalgo*
From Frugal to Reverse Innovation: A Knowledge-Driven Conceptual Framework for Upstream Diffusion

855 *Antonio Limongiello, Andrea Tomo*
Knowledge Transfer and Innovation Spillovers in Aerospace

864 *Giulia Padovani, Dario Barberini, Maria Giovina Pasca, Gabriella Arcese*
Climate Change Adaptation and Mitigation through Digital Technologies in the Extra Virgin Olive Oil Supply Chain

871 *Francesco Smaldone, Benedetta Esposito, Daniela Sica, Stefania Supino*
Mapping Circular Employability Skills in Agri-Food Value Chains from Large-Scale Labour-Market Data

878 *Paolo Canonico, Ernesto De Nito, Vincenza Esposito, Mario Pezzillo Iacono*
Evolving Competences through HR Projects: Knowledge Brokering Mechanisms in the SKiLL-PA Implementation Pathway

887 *Francesco Filippo, Francesco Paolo Lagrasta, Barbara Scozzi*
Evaluating the Acceptance of GenAI in Italian Academies of Fine Arts (AFAM)

897 *Antonio Lorena, Anna M. Correale, Alessia Talarico, Rocco Reina*
Artificial Intelligence and Soft Skills Development in Organizations: A Systematic Literature Review

907 *Michela Colacicco, Pierpaolo Pontrandolfo, Luigi Ranieri, Barbara Scozzi*
A Twin Transformation Maturity Model for Public Administrations: A Design Science Research Study

916 *Sini Tenhovuori, Nina Helander*
Experiences of Knowledge-based Human Resources Management in the Finnish Public Sector

924 *Francesca Lerario, Achille Claudio Garavelli*
Life Cycle Costing in Regenerative Systems: A Knowledge Integration Approach

930 *Umair Ali Khan, Dmitry Kudryavtsev, Akseli Leskinen, Janne Kauttonen, Jukka Remes*
From Unstructured Inputs to Structured Knowledge: A Modular LLM-Based Approach for Enterprise Knowledge Extraction

940 *Laura Iacovone*
Behaviour as Knowledge: Neuro-Immersive AI Assessment in HRM

949 *Jonna Käpylä*
Individual Intellectual Capital for Sustainability

958 *László Kállay*
Dynamic Capabilities and Economies of Scale: Rethinking Intelligent Knowledge in Sustainable Innovation

967 *Fabrizio Rossi, Alessandro Sicali*
Do Renewable Energy Communities Create Value? Evidence from a Systematic Literature Review

976 *Martina De Giovanni, Antonella Marsano, Mariangela Lazoi*
Rethinking Product Data Management Through Digital Product Passports: Data Governance for Global Supply Chains

985 *Sara Ianniello, Livio Cricelli, Serena Strazzullo*
Managing Sustainability in a Fashion Supply Chain: The Case of an Italian SME

995 *Tadej Crnjac, Robert Ojsteršek, Marta Aranyossy, Zlatko Nedelko*
Competence Model for Collaborative Workplaces in Industry 5.0: Initial Insight

1004 *Simona Mormile, Roberta Romano, Gabriella Piscopo, Paola Adinolfi*
Unveiling the Unseen: A Meta-Synthesis of Qualitative Informal Female Entrepreneurship Research

1012 *Stefania D'Aprile, Andrea Presciutti, Maria Zifaro*
Mapping Energy Innovation and ESG: A Bibliometric Review

1021 *Davide Madonna, Simone Malacaria, Andrea De Mauro*
Contextual–RFM (C-RFM): a Multimodal Framework for the Fusion of Transactional and Semantic Data Through LLMs for Interactive Customer Segmentation

1030 *Ferdinando Sasso, Andrea De Mauro*
Perturbation-First Robustness Evaluation of Automated Short Answer Grading

INDEX

1038 *Vincenzo Fiengo, Renato Passaro, Ivana Quinto*
Artificial Intelligence Applications in Biogas Systems: A Systematic Literature Review

1047 *Hayam Mohammed A Jabali, Ashish Malik, Philip J. Rosenberger III, Cathy Ying Xu*
Servant Leadership as a Relational Resource: Initial Insights from Qualitative Case Studies in Saudi Higher Education

1053 *Beren Çay, Mario Miozza, Andrea De Mauro*
Retrieval-Augmented Generation in Business Applications: A Systematic Literature Review and Taxonomy

1062 *Alessandro Margherita, Tina Comes, Gianluca Elia, Gianluca Solazzo, Istvan Hoffman, Jay Liebowitz*
AI-Enabled Coordination Framework to address Polycrisis: Human–Machine Collaboration and Regenerative Capacity

1071 *Mattia Vincenzo Olive, Maria Pinelli, Chiara Sgarbossa, Luca Gastaldi, Emanuele Lettieri*
Artificial Intelligence in Pharmaceutical R&D: Evidence on Trial

1077 *Roberta Romano, Simona Mormile, Gabriella Piscopo, Paola Adinolfi*
HRM Strategies for Diversity and Inclusion: A Bibliometric Analysis

1086 *Riccardo Censi, Paola Campana, Fulvio Schettino, Chiara De Pucchio, Valentina Ferrari*
From Quantitative ESG Data to Qualitative Social Insights: A Low-Code AI Agent for Decision Support

1094 *Marco Di Giovanni*
Can Ethics Become Isomorphic? Informational Power in GenAI Governance

1101 *Marta Menegoli, Luigi Jesus Basile, Roberto Mauriello*
A Regenerative Framework for Socio-Ecological Resilience: The Case of Armani's Apulia Regenerative Cotton Project

1107 *Mattia Arcuti, Vito Del Vecchio, Mariangela Lazoi, Giorgia Specchia*
Integrating ESG Principles into Production Processes for Shaping Sustainable Industry

1115 *Raimondo Fanale*
Explainable Artificial Intelligence as an Organisational Learning Lever: A Theoretical-Methodological Framework for Professional Upskilling in AI-Augmented Workplaces

1124 *Joachim Lorenz*
Tacit Knowledge Loss (TKL) from an Intelligent Knowledge-Based View (IKBV)

1131 *Marianna Mancino, Maria Zifaro, Giuseppe Ambrosio, Giovanni Spatola, Stefania D'Aprile*
Collective and Artificial Intelligence in the Challenge of Happy Degrowth

1141 *Gabriele Creta, Andrea De Mauro*
Explainable AI as Decision Support in Automated Assessment: An Exploratory Study

1149 *Nunzia Carbonara, Davide De Vita, Umberto Panniello, Elisabetta Vurchio*
How Healthcare Artificial Intelligence Startups Redesign Clinical Processes and Create Value in the Healthcare Ecosystem

1159 *Concetta Lucia Cristofaro, Marzia Ventura, Anna Maria Melina, Walter Vesperi*
Organizational Capabilities for Knowledge Ecosystems in Public Healthcare: Towards Intelligent and Sustainable Transformation

1166 *Stefania D'Aprile, Giovanni Spatola, Maria Zifaro, Giuseppe Ambrosio, Marianna Mancino*
Designing Hybrid Human AI Knowledge Ecosystems for Sustainable Value Creation

1175 *Alessandro Massaro, Giuseppe Loseto, Filippo Gramegna, Giovanni Schiuma, Francesco Santarsiero*
Transforming Cyberbullying Prevention Guidelines into Actionable Workflows through AI-Based Detection and BPMN Modeling

1182 *Pierluigi Mascaro*
Research Methods for Sustainable Organizations Using Intelligent Knowledge

1188 *Giovanni Maiuri, Domenico Celenza, Francesco Di Vona, Alessia Nardecchia*
Research on Megatrends and the Evolution of ESG Standard

INDEX

1196 *Domenico Celenza, Alessia Nardecchia, Francesco Di Vona, Marco Sebastiano*
The Experience of Error in Business Administration: The Case of Predictive Assessment in the Early Emergence of Corporate Distress

1204 *Domenico Celenza, Marco Sebastiano, Andrea Maturo, Antonella Fortuna*
Research on AI and Sustainability Ratings: Evidence from the Noverim Case

1212 *Alfonso Laudonia*
Intelligent Knowledge, Corporate Crisis and Criminal Liability: Incentive Structures for Sustainable Governance in the Italian Framework

1220 *Barbara Fasoli Braccini, Federico Gaeta*
From Cost to Social Capacity: Governing AI Released Time with STRR and S-STRR

1227 *Sara Gigli, Anna Turchetta*
How European Living Labs Contribute To the Creation of Ecosystems of Innovation: A Social Network Analysis (SNA) of Project-Based Collaboration

1234 *Fabrizio Rossi, Paola Rossi, Michele Bertoni, Domenico Celenza*
Legality Rating and Performance of Italian Companies: Face or Mask?

1242 *Giuseppe Ambrosio, Maria Zifaro, Federica Bandini, Stefania D'Aprile, Giovanni Spatola, Marianna Mancino*
Beyond Automation: Strategic Human-Machine Partnerships in Social Enterprise HRM. A Systematic Literature Review

1250 *Domenico Celenza, Marcella Caradonna, Francesca Francioli, Stefano Gazzani*
The Italian Crisis and Insolvency Code (CCII): Adequate Organisational Frameworks and AI Predictive Judgments

1257 *Antonio Lerro, Daniela Carlucci, Francesco Santarsiero, Giovanni Schiuma*
Leadership and Organizational Development in Remote Working Settings: Challenges, Opportunities and Empirical Evidences

1266 *Emilia Romeo, Nicola Capolupo*
When Does AI Become Knowledge? Organizational Conditions for Epistemic Learning in Public Administration

1273 *Giuseppe Bruno, Carmela Piccolo, Eduardo Pipicelli*
The Social Cost of Digital Banking: Branch Desertification and Inequalities in Access to Financial Services

1282 *Manuel Cavola, Patrice De Micco, Niccolò Parissi*
Assessing VSME Alignment in Voluntary Sustainability Reporting by SMEs: Evidence from the Italian Chemical Sector

1290 *Daniela Carlucci, Antonio Lerro, Francesco Santarsiero*
Driving Innovation and Digital Transformation in Cultural Tourism Organizations and Local Ecosystems: Final Insights of the T4Y Project

1298 *Benedetta Pini, Laura Monferdini, Barbara Bigliardi, Eleonora Bottani*
Artificial Intelligence in Supply Chains: A Systematic Review and Knowledge Mapping of Research Trends

1307 *Barbara Bigliardi, Virginia Dolci, Serena Filippelli, Benedetta Pini*
When Lead-User Signals Mislead: Entrepreneurial Sensemaking in Sport-Tech Innovation

1315 *Manfred Bornemann, Kay Alwert*
From Intellectual Capital Statements to AI-Enhanced Strategic Knowledge Management

1322 *Francesco Santarsiero, Giovanni Schiuma, Alessandro Massaro, Achille Claudio Garavelli*
Innovation Orchestration through Innovation Labs: Evidence from Intellimech

1330 *Lorella Cannavacciuolo, Carmine Passavanti, Pierluigi Rippa, Giustina Secundo*
Governing AI in Academic Innovation Ecosystems: A Process Model from Italian University Regulations

1338 *Gintare Kriaučiūnaitė-Lazauskienė*
Controversial Advertising: Multicultural and Generational Perspectives on Knowledge and Cultural Meaning-Making

1347 ASIA CHAPTER PAPERS

AI as a Driving Force for Sustainability, Resilience, Regeneration and Inclusiveness in the Context of Smart Cities

Maximilian Broglio
Mercatorum University, Rome, Italy

Abstract

AI has become a factor capable of disruptively shaping human existence, from society to corporate world, from economy to urban systems; however its focus on efficiency without regard for socio-ecological impacts, is now being strongly questioned; it is therefore essential to adopt a path of "humanization" of AI so as to allow it to evolve and align with the vision of the *Great Transition Network* (Raskin, 2021), which proposes a systemic transformation capable of not neglecting concepts such as social justice, socio-environmental sustainability, and human prosperity. AI must be considered a tool to achieve results in terms of sustainable development, which must always be the compass to follow in decision-making; this is where the concept of knowledge management evolves from the "mechanical" management of information to the concept of *intelligent knowledge* in which, through AI, systems are also capable of reasoning, automating complex processes, bridging the gap between raw data and human understanding, while minimizing any biases deriving from AI.

The aim of the paper is analysing, through a qualitative research approach (reviewing the most recent literature), the impact of AI as lever for sustainability, resilience, regeneration, and inclusiveness in the context of smart cities, starting from the assumption that cities must evolve, abandoning the application (without concern for impact) of efficiency and intensification, embracing instead the idea of creating ecosystems oriented towards socio-environmental sustainability, cooperation and participation.

After the specification of the objectives of a more "humanized" AI along with the areas of intervention where AI can effectively operate and the risks associated with its shortsighted implementation, case studies of cities that have successfully implemented these innovative technological strategies are presented (selected according to relevance, transformative potential, and replicability with no restraints in terms of geographical boundaries).

Findings support the view that a sustainable implementation of AI can lead to sustainability, resilience, regeneration, and inclusiveness in the context of smart cities, promoting social and environmental equity (lower energy consumption, optimized resource management, better climate resilience, social inclusion). Examples in this sense can be Masdar City (UAE), where traffic congestion is reduced through AI systems capable of calibrating routes based on real-time weather and traffic information, Fukuoka (Japan) that, through the exploitation of AI-based sensors in its streetlights system, is saving up to 25% on electricity, Los Angeles (USA) that has applied predictive analytics to prevent homelessness by proactively identifying individuals at heightened risk of losing their home.

Keywords: AI; smart cities; sustainability; resilience; inclusiveness.

Paper type: Academic Research Paper

1 Introduction

Artificial intelligence is a multidisciplinary field of computer science characterized by a set of technologies that allow a machine to emulate human capabilities through algorithms that learn from available information and/or information collected directly by the machine to improve its performance over time; there are several types, including *Machine Learning*, which is characterized by algorithms capable of automatically learning from existing data and making predictions, *Deep Learning*, which uses deep artificial neural networks to learn complex representations from data (e.g. images and video), *Generative Artificial Intelligence*, in which algorithms, trained on large amounts of data, can generate new outputs similar to the original data and *Agentic AI* designed to pursue specific goals autonomously with capabilities to make decisions and take actions based on feedbacks and increasingly used in computer engineering, in robotics or in decision-support systems (Casper et al., 2025); noteworthy, dealing with smart cities, is the concept of *Edge AI*: it combines AI with edge computing, enabling machine learning algorithms and data analysis to be run directly on local devices such as cameras and sensors enabling real-time decisions, saving bandwidth and improving privacy (data remains on the device).

AI requires to be somehow "humanized" in order not to be exploited in a shortsighted race for efficiency devoid of any socio-ecological implications. In this context, *intelligent knowledge* was born, an evolution of the

traditional concept of knowledge management, leveraging the potential of AI, going beyond simple information retrieval to enable intelligent decisions and automate complex processes, bridging the gap between raw data and human understanding. Starting from the fact that knowledge management is linked to the mechanical management of information, while intelligence is linked to the ability to analyse it, *intelligent knowledge* represents a fusion, creating solutions capable of reasoning and deduction, reducing potential AI biases.

Analysing the impact of AI in urban context, a paradigm shift must be implemented with sustainability as guiding principle: AI must not be seen as a mere efficiency tool, but as a key tool for addressing all ESG aspects that holistically define the concept of sustainability. Cities are transitioning from "smart cities" to "cognitive cities", not because the underlying technology is different, but because the governance paradigm must evolve from a "technology-first" to a "human-first" approach.

The purpose of the paper is analysing the impact of AI as lever for sustainability, resilience, regeneration, and inclusiveness in the context of smart cities; these ones are cities capable of leveraging digital technologies such as IoT, AI, and Big Data to evolve their services and ecosystems, for example in the areas of urban planning, transportation, energy management, and waste management. Their goal is improving the overall quality of life of their inhabitants, socio-environmental sustainability, and operational efficiency, actively engaging citizens and businesses for sustainable development.

Urban contexts must evolve as fertile grounds for experimentation, abandoning the application (without regard to impacts) of efficiency and intensification, embracing instead the idea of creating ecosystems oriented towards socio-ecological sustainability, regeneration, resilience, and inclusiveness, where AI can also promote responsibility, cooperation, participation, sufficiency, and sharing, thus paving the way for more inclusive innovation (Zhang et al., 2025).

This paper analyses areas of intervention where AI can operate to achieve the sustainability goals promoted by the *Great Transition Network* and discusses case studies of smart cities that have successfully adopted and implemented these innovative technological strategies in a sustainable manner.

The ultimate goal of the paper is to act as a driving force for cities that are still hesitant and to demonstrate how a conscious and calibrated use of these digital technologies can truly contribute to the creation of sustainable, resilient, regenerative, and inclusive smart cities. The aim is triggering a process of "constructive emulation" in which cities, leveraging existing experiences and emerging opportunities, can provide a decisive boost to the evolutionary path toward more sustainable and contemporary cities.

2 Methodology

This study employs a qualitative research approach grounded in the Multilevel Perspective (Geels, 2002) with the aim of achieving tangible findings regarding the contribution of a calibrated and conscientious use of AI to the sustainable development of smart cities.

The research was conducted in two phases: the first involved a comprehensive review of the most recent technical literature to analyse the specific characteristics of a sustainable AI-approach in urban settings, defining objectives, areas of intervention, and potential negative impacts in order to provide an in-depth and up-to-date overview; the second involved a case studies review of cities that have successfully implemented innovative digital solutions, in order not only to provide a detailed overview, but also to act as a sounding board for other cities that have yet to take this step (case studies were selected based on their relevance, specificity, transformative potential, and replicability in the context of smart cities with no restraints of specific geographical boundaries; the only constraint is that they were recent and successful in order to represent a benchmark).

In terms of research framework, this work is connected to two research streams: AI, which is monopolizing interest in a wide variety of research fields, and sustainability, a really hot topic, as confirmed by the number of specific academic publications; the study stands out in a context of hybridization, considering concepts that synergistically intertwine between the two aforementioned streams. This research, like all sustainability research, is also part of the ESG framework. The specific topic of smart cities is gaining increasing attention in the research world as it relates to a topic that is proving to be particularly crucial for the sustainable and inclusive evolution of global urban systems, which must necessarily curb the negative spiral associated with the global context in which they exist (anthropogenic climate-altering emissions, indiscriminate use of non-renewable resources, biogeochemical pollution, social injustices, demographic increase, etc.).

3 Findings

3.1 Goals, areas of intervention, and risks of AI in the context of smart cities

Findings highlight how the careful use of AI in smart cities can effectively address several crucial issues, such as sustainability, resilience, regeneration, and inclusiveness without neglecting that an inappropriate use can lead instead to significant risks.

Leveraging these innovative digital technologies, capable of exploiting real-time data from sensors and IoT, and integrating predictive simulations for low-impact planning, can promote the optimization of urban resources and services, predict needs, and support informed and timely decisions reducing socio-environmental impacts, shaping cities that increasingly transform into intelligent, resilient, and adaptive ecosystems.

Four areas of intervention are identified where AI can operate in the smart cities' context: sustainability, resilience, regeneration, and inclusiveness.

In the area of sustainability, AI can have a significant impact in the environmental sphere: AI can lead to the optimization of resources and energy consumption (e.g. in the context of the smart grid[1]) and to a more efficient management of vehicular traffic reducing the impact in terms of emissions, it can help make water and waste management services more efficient, reducing waste and encouraging a transition to circular models and can support urban agriculture; within sustainability area, interventions can be categorized into four application sectors: energy management, mobility, waste/water management, and urban planning. In terms of energy management, AI can intervene through intelligent regulation of smart grids and buildings, to reduce consumption through adaptive air conditioning and lighting systems and allowing the integration of renewable energy sources (e.g. the Hudson Yards in New York); in terms of mobility, AI enables smarter traffic management through the use of predictive algorithms capable of managing traffic flows: in Masdar City (UAE), traffic congestion is reduced through AI systems capable of calibrating routes based on real-time weather and traffic information while "RoadSafeAI", a predictive model developed by the Polytechnic University of Milan and Unipoltec, exploits convolutional neural networks to analyse road images and telematics data, identifying high-risk areas for accidents with an accuracy of over 95%. For the environmental sector, AI can manage the monitoring of waste collection and recycling cycles through predictive models, optimizing routes, reducing transportation-related emissions, and promoting effective resource management. Similarly, through the use of predictive models and intelligent monitoring, it can ensure a reduction in water waste and a general systemic improvement in the integrated water service. For the urban planning sector, AI can support eco-design based on climate data and demographic assessments to guide adaptive projects; AI (in planning) can help identify pollutants and consequently plan green spaces to improve air quality and biodiversity (e.g. AI-enabled BIM can simulate the impacts of materials/buildings), anticipate sustainable expansions with tools such as those from IBM and SEforALL[2], and promote the circular economy by tracking resources.

In the area of resilience, AI works primarily in an anticipatory manner through predictive algorithms that, by processing large volumes of real time data from IoT sensors and integrating technologies such as digital twins, allow for the prevention of various types of emergencies: environmental emergencies such as earthquakes[3], heat waves[4], landslides and floods[5], system emergencies as blackouts or emergencies related to the digital resilience (the ability to anticipate, resist, respond and recover quickly from interruptions, cyberattacks, IT failures, ensuring operational continuity and protecting data, services, processes and technologies keeping critical functions active before, during and after an incident); in addition to forecasting, there is also monitoring: the use of AI-based automatic alert systems capable of monitoring traffic, fires, or floods and providing timely notifications to citizens and authorities significantly increases urban resilience. The impact of AI on urban resilience during emergency management must be also considered: AI optimizes evacuation routes and resources allocation, improving safety and response capacity, also leading to an accurate post-event damage assessment to prioritize interventions. It is fundamental to strengthen data-informed governance by enabling policymakers to anticipate problems, model scenarios and design adaptive responses to increase the overall resilience.

Another area of intervention is urban regeneration: in this case, generative AI models can directly influence the design of new spaces, promoting accessibility and sustainability in design choices (for example, supporting

---

[1] which in turn, by facilitating the integration of renewable sources and the management of distributed production, allows for the reduction of climate-altering emissions.

[2] SEforALL (Sustainable Energy for All) is an international organization working in partnership with the United Nations, government leaders, the private sector, financial institutions, and civil society to promote faster and more concrete action to achieve Sustainable Development Goal 7.

[3] Smart buildings with AI sensors can assess real time safety in case of earthquake.

[4] Vulnerability maps can identify at-risk areas and vulnerable patients.

[5] For example, Lisbon is leveraging AI-driven digital-twin modelling to simulate flood scenarios and plan adaptive infrastructure responses.

energy communities); through the analysis of large volumes of socio-environmental data and the use of simulations of intervention scenarios to predict impacts on different targets, the most suitable areas (degraded and problematic areas deserving of higher priority) can be selected and the most sustainable solutions can be implemented.

The final area of intervention examined concerns inclusiveness (linked to the "S" of ESG); AI works primarily on three aspects: digital accessibility, social inclusion, and inclusive design. In terms of digital accessibility, AI offers solutions for people with disabilities, such as voice recognition systems, AI-enhanced screen readers, automatic subtitles, and interface customization; it also enables access to services such as transportation or healthcare for non-native speakers through multilingual chatbots and virtual assistants. In terms of social inclusion, by monitoring inequalities through predictive analytics, enables interventions on digital literacy to prevent exclusion and supports inclusive work adapting tasks to the specific disability. AI also promotes inclusion and social sharing of benefits through integration of ethics and privacy to ensure fairness. For inclusive design, generative AI works on simulating accessible spaces, seeking to meet needs of people with disabilities, while also leveraging citizen participation (e.g. digital twins allow for exploiting citizen feedbacks on future projects to increase transparency[6] and engagement).

Social inclusion is one of the most insidious aspects dealing with the link between smart cities and AI: there is the risk, in administrators' strategic choices, of prioritizing issues like urban efficiency and economic growth, thus risking exacerbating social inequalities. These considerations are supported by a recent study (Ntanda and Carolissen, 2025), which highlights a complex relationship between smart cities development and social inequalities, characterized by the risk of amplifying existing disparities if not managed inclusively and requiring robust community involvement and adaptation of innovations to the local context.

AI is characterised by possible unsustainable aspects that must be taken into account analysing its relationship with smart cities: for Vinuesa et al. (2020), AI has the potential to have negative impacts on sustainability[7]; three types of risks can be considered: ecological risks (AI footprint)[8], socio-economic and ethical risks[9], risks related to data/algorithms[10]; in relation to this, the concept of the "green AI paradox" can be considered, i.e. the risk that the AI's environmental costs outweigh the benefits of AI-implemented green solutions: this aspect must certainly be addressed dealing with the aforementioned risks[11]. Furthermore, it is impossible to ignore the Jevons Paradox[12], which posits that technological improvements capable of increasing the efficiency can lead to an increase in the total consumption of a resource.

Moreover, the success of smart cities is amplified by the combination 'open data-AI' in which open data can provide a large amount of data for AI systems to function effectively[13].

3.2 Case studies

It is compelling to analyse case studies of AI application to support sustainable and inclusive smart cities in order to evaluate types of intervention, application rates and scalability with the ultimate goal of inspiring constructive emulation by which cities can embark on a path toward sustainability and inclusiveness.

In terms of characteristics, adoption rates and types of AI use, interesting results can be deduced from the study "AI-powered cities of the future" (2025)[14] based on the AI plans, investments, and practices of 250 cities in 78 countries; first of all what can be deduced is that 56% of surveyed cities actively use AI with a projected adoption rate up to 83%[15] in the next-three-years period. In addition, the study shows that in terms of smart

---

[6] An example in which there is the exploitation of real-time data to support transparent decision-making and collaborative urban planning, offering a practical example of how technology can enable sustainable and inclusive urban development is the city of Rotterdam (The Netherlands).

[7] according to authors 59 of the 169 sustainable development goals could be negatively impacted.

[8] For example, in terms of energy consumption (training AI models, especially large ones like deep learning models, requires a significant amount of energy), water consumption (some data centres use water cooling systems, which can consume significant amounts of fresh water), e-waste management (the rapid obsolescence of hardware technologies used for AI contributes to the growing amount of e-waste). All these factors can be summarized in the concept of digital pollution.

[9] For example, in terms of potential job losses (AI-powered automation can replace human jobs, especially in some sectors), technology dependence (the increased use of AI in various areas of daily life can lead to an overdependence on technology), security (AI can be used for malicious purposes), unequal access to technology (risk that AI will become the privilege of a few, accentuating inequalities).

[10] For example, in terms of "black box" responsibility, erosion of digital trust, systemic digital bias (AI algorithms can reflect and amplify biases present in the data they are trained on thus leading to discriminatory decisions).

[11] It should also be emphasized that achieving this goal also requires governance and international cooperation frameworks, financing and incentive policies for green AI, international collaboration, open-source platforms, and training programs.

[12] Named after the British economist and logician William Stanley Jevons who lived during the nineteenth century. Linked to the concept that greater efficiency reduces costs and stimulates greater demand, expanding overall use.

[13] an example is the Irish city of Dublin which demonstrated how open data help smart cities using digital twin technology for data-driven decision-making.

[14] Deloitte, ServiceNow, NVIDIA, ThoughtLab (2025) AI-powered cities of the future.

[15] this data is interesting due to the fact that, internationally, comparable data on levels of AI adoption across cities remain rare. In terms of adoption rate, it is compelling to underline that is in progress a faster than ever AI adoption, particularly in city administration and mobility, and in cities like Seoul (Korea) that have pioneered the integration of AI into urban infrastructure (The Smart City Index 2024, "AI Transforming Cities: Insights from Professor Jung-hoon Lee on the Smart Cities Index 2024". Based on the analysis of over 5000 smart city services, infrastructure and initiatives across 50 cities).

cities development, large metropolitan areas steal the spotlight[16] (in Europe cities considered pioneers in the most advanced implementation of AI are among the others Amsterdam, Barcelona, Berlin, Dublin, Helsinki, Madrid, Paris, Vienna while, outside of Europe, are New York, Boston, Chicago, Los Angeles, Toronto, Montreal, Sao Paulo, Istanbul, Beijing, Hong Kong, Melbourne, Seoul, Sydney and Tokyo) certainly related to greater economic possibilities, opportunities for research and development, possibility to build a modern IT foundation, greater ease in developing AI skills, talent, processes, possibility to combine AI with other technologies and greater propensity for innovation and change. It is interesting to underline that in this case ten areas of AI intervention in urban context are considered: traffic management, public safety, energy management, customer safety, predictive management, waste management, data analysis, urban planning, cybersecurity, fraud detection.

It is now compelling to review some examples[17] of AI applications at the service of sustainability, resilience, regeneration and inclusiveness so that they can serve as benchmarks for urban contexts that want to approach the world of AI to begin to accelerate their path toward sustainable development.

Fukuoka (Japan) through the exploitation of AI-based sensors in its streetlights system is saving up to 25% on electricity, Seattle (USA) is experimenting with an intelligent AI-based parking management system able to cut down on parking durations thus enhancing air quality, Shijiazhuang (China) is using AI in waste management in order to increase efficiency and recycling, Madrid (Spain), in the context of the EU-funded project "AMIGOS", uses AI-driven smart cameras and sensors to monitor city spaces in Las Rozas, providing real-time data helping city planners make informed decisions (a similar initiative, "ELABORATOR", is being implemented in the French district of Issy-les-Moulineaux, and is characterized by the usage of AI to automatically manage intersections, creating more space for cyclists and alerting drivers to their presence based on traffic conditions). Singapore, in its Changi Airport, uses AI to screen and sort baggage, as well as to power behaviour recognition technology for immigration clearance enhancing border security, Jaipur (India) implements AI to effectively manage water harvesting and environmental impact assessment reports in industrial areas to mitigate pollution, Dubai (UAE) is using AI to improve daily life with customer service chatbots (some upgraded to use generative AI, including those for the Roads and Transport Authority, and the Dubai Health Authority), and for mobility with an AI-powered traffic management system, San Francisco (USA) is exploring the use of Agentic AI, enabling machines to take decisions and actions without direct human intervention, through the usage of self-driving taxis, Raleigh (USA) uses a full digital twin of the city, paired with an AI-powered microclimate modelling system, for more intelligent and resilient urban development and also uses AI to forecast flooding in stormwater systems, based on factors such as age and condition of infrastructure, soil materials, and previous failures. Tokyo (Japan) is using a comprehensive disaster preparedness system consisting of a network of AI sensors combined with predictive modelling, simulation tools, and community-based disaster management initiatives, Yizhuang-Beijing (China) is using AI in the healthcare sector developing a cloud-based infrastructure that connects medical institutions across Beijing (including intelligent triage, pre-diagnosis consultations, and appointment scheduling), Sao Paulo (Brazil) has developed an AI predictive maintenance tool that can anticipate faults in its metro and has developed a mobility-as-a-service app to facilitate the transportation through the city for residents, Venice (Italy), to address overtourism, has implemented an AI-enabled smart control room, which gathers real-time data from existing digital infrastructure assets, including intelligent sensors and video cameras, to optimize network functionality. Melbourne (Australia) employs AI and uses sensors to monitor illegal waste dumping behaviours and to track trees in the city and make plans for future planting, Barcelona (Spain) exploits an AI-powered ventilation control platform called "RESPIRA" to address challenges of particulate matter pollution, thermal stress, and ventilation inefficiencies, Los Angeles (USA) has applied predictive analytics to prevent homelessness by proactively identifying individuals at heightened risk of losing their home, Eslöv (Sweden) uses an AI-assisted planning and design platform called "Hektar" characterized by generative design and data-driven analysis to support early-stage urban planning and zoning optimisation, Praga (Czech Republic) to improve the accessibility of urban services uses AI chatbot, Montevideo (Uruguay) has employed AI and predictive analytics to identify neighbourhoods most vulnerable to flooding, poverty, and service deprivation (to better prioritise social and infrastructural investments). Maastricht (The Netherlands) exploits soil moisture sensors linked to AI-based analytics to enable tree managers to respond to real-time alerts during heatwaves, Sydney (Australia) uses Google's Environmental Insights Explorer (EIE) "Tree Canopy", which applies machine learning to aerial imagery, to visualise canopy gaps at neighbourhood scale and plan where trees are most needed for heat-vulnerable areas, Bogotá (Colombia) uses AI-based chatbot "Candelaria" to strengthen perceptions of safety, improve

[16] however, small cities such as Drancy in France, prove that size is no barrier to innovation even if obviously the starting conditions are not the same.

[17] Found, among the others, in Deloitte, ServiceNow, NVIDIA, ThoughtLab (2025) AI-powered cities of the future and in OECD (2025) Artificial Intelligence for advancing smart cities, 5th OECD Roundtable on smart cities and Inclusive Growth.

tourist information and increase visitor engagement, Jammerbugt (Denmark) uses an AI-based early warning system for flooding integrating satellite imagery, weather forecasts and hydrological models to predict floods up to 48 hours in advance, Seoul (South Korea) has applied computer vision to the city's extensive CCTV network (more than 110.000 cameras) and launched an AI-enabled CCTV initiative to strengthen prevention, improve response times and build public trust in safety services by recognising behaviours such as wandering, aggression, crowd density.

4 Conclusions

Considerations thus far support the view that a correct and sustainable use of AI can lead to sustainability, resilience, regeneration, and inclusiveness in the context of smart cities. AI can promote social and environmental equity, leading to the creation of more liveable cities with lower energy consumption, optimized resource management, reduced waste, greater climate resilience, social inclusion.

It should be emphasized that the paradigm must evolve from a "technology-first" approach focused more on efficiency and intensification to a "human-first" approach that places human well-being at the centre, accounting for all ESG aspects holistically and in which technology is a tool at the service of social inclusion, environmental sustainability and economic development (where attention is paid not to reinforce digital divides and not to fall in excessive monitoring and privacy violations). It must also be noted that the progressive implementation of AI in smart cities must inevitably follow an incremental approach to allow cities to gradually reach maturity without simultaneously losing public trust: this can mean starting with strategically selected operational pilot projects that offer immediate value and then gradually expanding its use.

Another key concept is that of "constructive emulation": successful sustainable AI implementation can undoubtedly serve as a boost for other urban systems that have not yet integrated AI systems into their strategies, thus triggering emulation that can lead to a globally broadened impact in terms of sustainability, resilience, and inclusiveness.

Systemic scaling of AI is most feasible when national regulation and experimentation instruments are coordinated under a unified architecture: a top-down commitment is required, where laws are harmonized, coherent and effective and are capable of increasing citizens' trust (without clipping the wings of the innovative potential of AI); laws and plans as EU's AI Act (2024)[18], China's Algorithmic Recommendation Regulation (2022)[19] or the open AI registers[20] aimed to be responsible and transparent about the use of algorithms and AI, are fundamental to send a strong signal of transparency and trust.

In terms of limits, this paper does not take into account barriers that hinder AI use (implementation and maintenance costs, required skills, resistance to change)[21] and regulatory frameworks (to fully understand the frameworks within which AI-based strategies can flourish without incurring the risks associated with its unsustainable and erroneous use).

In terms of smart cities future developments, according to Zain (2025) one option is the *predictive city* where AI continuously analyses data from sensors and digital twins to anticipate needs and optimise services while another option (*participatory intelligent city*) may be characterized by an AI designed to increase collective intelligence where data and algorithms are open and co-created with citizens.

Research must seek to increasingly explore the connection between AI and sustainability in the context of smart cities through the analysis of innovative solutions, with the hope that the number of cities that can become benchmarks can further expand, as this would indicate that the process of constructive emulation is fully underway.

5 Ethics declaration

Ethical clearance was not required for the research.

6 AI declaration

No artificial intelligence tools were used to generate content or conduct analyses.

---

[18] *Regulatory framework to ensure the safe and ethical use of AI.*

[19] *Regulation that requires transparency, fairness and prohibits AI applications that promote illegal content, manipulate opinions, or false accounts.*

[20] *The AI register (e.g. prepared by the cities of Helsinki and Amsterdam) describes for example which datasets are used for training purposes, how are the algorithms and models assessed for potential bias or risks, in which ways do humans use the AI services, etc.*

[21] *that can be addressed through public funding, public-private consortia, training, and specific education.*

References

Adeyemo Adepoju, P. et al. (2025) "Smart Cities and Internet of Things (IoT): A Review of Emerging Technologies and Challenges", *International Journal of Research and Innovation in Social Science*, 2025, vol. 9, issue 1, 1536-1549.

Alhasnawi, B. N. (2025) "The rising, applications, challenges, and future prospects of energy in smart grids and smart cities systems", *Energy Conversion and Management: X* Volume 27, July 2025, 101162.

Alizadeh, H. et al. (2023) "Toward a societal smart city: Clarifying the social justice dimension of smart cities", *Sustainable Cities and Society*, 95 (2023), Article 104612.

Al-Raeei, M. (2024) "The smart future for sustainable development: Artificial intelligence solutions for sustainable urbanization", *Sustain. Dev*. 2024, 33, 508–517

Badidi, E. (2022) "Edge AI and Blockchain for Smart Sustainable Cities: Promise and Potential", *Sustainability* 2022, 14, 7609.

Bansal, C. (2025) "AI ethics and sustainability: Accelerating paradigm shifts toward sustainable development", *Journal of Strategy & Innovation*, Volume 36, Issue 1, May 2025, 200537.

Belaid, F. et al. (2025) "Can Smart Cities Be Sustainable? An Emerging Field of Research", *Journal of Urban Technology*, 32(1), 27–48.

Chen, S. (2025) "How much energy will AI really consume? The good, the bad and the unknown", *Nature* 2025, 639, 22–24.

Clark, J. (2020) *Uneven innovation: The work of smart cities*. Columbia University Press.

Deloitte, ServiceNow, NVIDIA, ThoughtLab (2025) *AI-powered cities of the future*.

De Vries, A. (2023) "The growing energy footprint of artificial intelligence", *Joule* 7, 2191–2194.

Elassy, M. (2024) "Intelligent transportation systems for sustainable smart cities", *Transportation Engineering*, Volume 16, June 2024, 100252

Evans, J. et al. (2019) "Smart and sustainable cities? Pipedreams, practicalities and possibilities", *Local Environment*, 24 (7) (2019), pp. 557-564.

Goh, H.H., Vinuesa, R. (2021) "Regulating artificial-intelligence applications to achieve the sustainable development goals", *Discover Sustainability*, 2 (1) (2021).

Guan, L. (2024) "Reaching carbon neutrality requires energy-efficient training of AI", *Nature* 626, 33–33.

Hardi, I. et al. (2025) "Inclusive innovation and green growth in advanced economies", *Sustainable Futures*, Volume 9, June 2025, 100540.

Jia, L. (2023) "Which type of innovation is more conducive to inclusive green growth: independent innovation or imitation innovation?", *J. Clean. Prod*., 406 (2023).

Kalkanci, B. (2019) "The role of inclusive innovation in promoting social sustainability", *Prod. Oper. Manag*., 28 (2019), pp. 2960-2982.

Khlewee, I. K. et al. (2025) "AI in Smart Cities: A Review of Urban Data Processing, Prediction, and Optimization Techniques", *ESTIDAMAA*, 2025, 29-38.

Kolotouchkina, O. et al. (2024) "Smart cities, digital inequalities, and the challenge of inclusion", *Smart Cities*, 7 (6) (2024), pp. 3355-3370.

Lowe, N., Feldman, M. P., Zoller, T.D. (2021) "Inclusive innovation: Defining and measuring inclusiveness in innovation districts". *Journal of Urban Affairs*, 43(5), 736–738.

Luo, X. Huang, R. (2024) "How to achieve sustainable development of smart city: A comprehensive bibliometric analysis", *Clean Technologies and Environmental Policy* (2024).

Morales, D. et al. (2025) "Defining inclusive innovation: challenges and lessons for innovation districts", *Regional Studies, Regional Science*, 12:1, 740-753.

Nazmul Haque, Md et al. (2025), "A systematic review of research on just, equitable, responsible, and inclusive smart cities", *Technology in Society*, Volume 83, December 2025, 103050.

Nesmachnow, S. et al. (2025) "A Literature Review of Recent Advances on Innovative Computational Tools for Waste Management in Smart Cities", *Urban Sci*. 2025, 9, 16.

Ntanda, A., Carolissen, R. (2025) "Technology's dual role in smart cities and social equality: A systematic literature", *Journal of Local Government Research and Innovation*.

OECD (2025) *Artificial Intelligence for advancing smart cities*, 5th OECD Roundtable on smart cities and Inclusive Growth.

Pandya, S. et al. (2023) "Federated learning for smart cities: A comprehensive survey", *Sustainable Energy Technologies and Assessments*, Volume 55, February 2023, 102987.

Patel, Y. et al. (2019) "Social implications of smart cities", *Procedia Computer Science*, 155 (2019), pp. 692-697.

Qanazi, S. et al. (2025) "Integrating social dimensions into urban digital twins: A review and proposed framework for social digital twins", *Smart Cities* 2025, 8, 23.

Rejeb, A. et al. (2022) "The big picture on the internet of things and the smart city: A review of what we know and what we need to know", *Internet Things* 2022, 19, 100565.

Sacoto-Cabrera, E. J. (2025) "IoT, AI, and Digital Twins in Smart Cities: A Systematic Review for a Thematic Mapping and Research Agenda", *Smart Cities* 2025, 8(5), 175.

Santos, S. et al. (2024) "Artificial Intelligence in Sustainable Smart Cities: A Systematic Study on Applications, Benefits, Challenges, and Solutions". *Sustainability* 2024, 1, 644–655.

Shao, J. et al. (2025) "Sustainable development strategies for Smart Cities: Review and development framework", *Cities* Volume 158, March 2025, 105663.

Sharma, A. et al. (2025) "Urban energy transition in smart cities: A comprehensive review of sustainability and innovation", *Sustainable Futures*, Volume 10, December 2025, 100940.

The Smart City Journal (2024) “AI Transforming Cities: Insights from Professor Jung-hoon Lee on the Smart Cities Index 2024”.
Toderas, M. (2025) “Artificial Intelligence for Sustainability: A Systematic Review and Critical Analysis of AI Applications, Challenges, and Future Directions”, *Sustainability*. 2025; 17(17):8049.
Trencher, G. (2019) “Towards the smart city 2.0: Empirical evidence of using smartness as a tool for tackling social challenges”, *Technological Forecasting and Social Change*, 142 (2019), pp. 117-128,
Tura, N. et al. (2022) “Sustainability-oriented innovations in smart cities: A systematic review and emerging themes”, *Cities*, 126 (2022), Article 103716.
Van Wynsberghe, A. (2021) “Sustainable AI: AI for sustainability and the sustainability of AI”, *AI and Ethics*, 1 (3) (2021), pp. 213-218.
Velaga, K.S. et al. (2025) “Edge AI for Smart Cities: Foundations, Challenges, and Opportunities”, *Smart Cities* 2025, 8, 211.
Vinuesa, R. et al. (2020) “The role of artificial intelligence in achieving the Sustainable Development Goals”, *Nature communications*, 11(1), 1-10.
Yadav, V., Yadav, N. (2024) “Beyond Sustainability, Toward Resilience, and Regeneration: An Integrative Framework for Archetypes of Regenerative Innovation”, *Glob J Flex Syst Manag* 25, 849–879.

Artificial Intelligence and Sustainability Management as Innovation Strategies for Sustainable Competitiveness: Evidence from Innovative SMEs

Giovanna Ferraro[1], Ivana Quinto[2], Giuseppe Scandurra[2], Antonio Thomas[2]
[1] University of Rome Tor Vergata, Rome, Italy
[2] Parthenope University of Naples, Naples, Italy

Abstract

In recent years, environmental, social, and governance (ESG) practices have progressively become a central component of corporate strategies and investment decisions. Although ESG is often associated with corporate social responsibility (CSR), the two concepts differ in important ways. CSR typically refers to voluntary initiatives oriented toward responsible business conduct, whereas ESG is based on standardized and measurable indicators that enable stakeholders to assess companies' sustainability performance in a more systematic and comparable manner. The growing relevance respect to ESG practices is driven by increasing societal awareness of sustainability challenges, stronger stakeholder expectations, and the evolution of regulatory frameworks aimed at promoting sustainable development However, the mechanisms that support ESG implementation among small and medium-sized enterprises (SMEs), as well as the potential competitive benefits associated with these practices, remain only partially explored. This study investigates the role of artificial intelligence (AI) and sustainability management (SM) as key drivers of ESG adoption and competitive perspectives in innovative SMEs. The empirical analysis is conducted using Partial Least Squares Structural Equation Modelling (PLS-SEM) on a sample of 163 Italian innovative SMEs located in Lazio Region. Data were collected through a structured online questionnaire administered to company representatives involved in strategic and innovation-related decision-making processes. The results show that both AI and SM significantly contribute to the adoption of ESG practices across environmental, social, and governance dimensions. However, only AI exerts a direct positive effect on competitive perspectives, while SM influences competitiveness indirectly through ESG practices. Furthermore, the ESG dimensions do not equally affect competitive outcomes. The environmental and governance dimensions positively influence competitive perspectives, whereas the social dimension does not exhibit a statistically significant effect. These results provide new insights into the mechanisms through which technological and managerial factors support sustainability-oriented strategies and competitiveness in SMEs, highlighting the importance of integrating digital capabilities and structured sustainability practices.

Keywords: Artificial intelligence; Sustainability management; ESG practices; SMEs competitiveness
Paper type: Academic Research Paper

1 Introduction

The adoption of ESG practices is increasingly relevant for companies of all sizes. Despite the increasing diffusion of ESG frameworks, their adoption among SMEs remains, in many cases, limited. SMEs frequently face structural constraints, including limited financial resources, lack of specialized managerial competencies, and difficulties in accessing tools for sustainability measurement and reporting (Rubino et al., 2024). These barriers contribute to a fragmented implementation of ESG practices and make it difficult to clearly identify the drivers that support their adoption. The existing literature has proposed several theoretical perspectives to explain companies' engagement in sustainability and ESG practices, including stakeholder theory, institutional theory, and legitimacy theory. These approaches mainly emphasize the role of external pressures, regulatory requirements, and the need for firms to align with stakeholder expectations. However, the internal mechanisms through which SMEs translate these pressures into concrete sustainability practices are still not fully understood. Recent studies suggest that both technological and managerial factors can play a crucial role in supporting ESG adoption. On the one hand, artificial intelligence (AI) enables companies to analyze large volumes of data, improve decision-making processes, and enhance the monitoring of environmental and social performance (Cheng et al., 2024; Tutore et al., 2024). On the other hand, sustainability management (SM) provides the organizational structures, routines, and policies necessary to embed sustainability principles into business strategies and operational activities (Lei and Yu, 2023). Although the role of digital technologies and sustainability-oriented managerial practices has received increasing attention, empirical evidence on how these factors jointly contribute to ESG adoption and influence companies' competitive perspectives remains limited. This gap is particularly relevant in the context of SMEs, which represent a fundamental component of economic systems and innovation ecosystems but often struggle to translate sustainability initiatives into tangible competitive outcomes. This study investigates the role of AI and SM as key drivers of ESG adoption and

competitive perspectives among innovative SMEs. The research focuses on understanding whether and how these factors support the implementation of ESG practices and whether ESG adoption contributes to strengthening companies' competitive positioning.

The study addresses the following research questions:

- *RQ1: How do AI and SM jointly influence the adoption of ESG practices in SMEs?*
- *RQ2: To what extent does the adoption of ESG practices contribute to enhancing SMEs' competitive perspectives?*

To answer these questions, the study develops a conceptual model and empirically tests it using Partial Least Squares Structural Equation Modelling (PLS-SEM) on a sample of 163 innovative Italian SMEs. These firms are characterized by strong technological orientation, high levels of human capital, and systematic investments in research and development, making them a particularly relevant context for analyzing the interplay between technological capabilities, sustainability practices, and competitiveness.

The paper is structured as follows. Section 2 presents the theoretical background and research hypotheses. Section 3 describes the methodology and data collection process. Section 4 reports the empirical results, Sections 5 and 6 discuss the findings and present the conclusions.

2 Theoretical Background

2.1 ESG Practices and SMEs

ESG practices have emerged as one of the most widely used frameworks for evaluating corporate sustainability performance. Through ESG indicators, stakeholders are able to assess companies' environmental impact, social responsibility, and governance quality using standardized and comparable criteria. Over time, ESG metrics have increasingly influenced corporate strategies, investment decisions, and stakeholder expectations (Sánchez-Torné et al., 2020; Jiang et al., 2024). Although ESG frameworks were initially developed for large corporations, growing attention has been directed toward their adoption among SMEs. SMEs represent a key component of national economies and innovation systems, yet their approach to sustainability often differs from that of larger firms. Due to their limited financial resources, less formalized organizational structures, and restricted access to specialized expertise, SMEs frequently encounter significant challenges in implementing ESG practices (Rubino et al., 2024). Despite these constraints, existing studies suggest that ESG adoption may generate several benefits for SMEs, including enhanced corporate reputation, improved stakeholder relationships, better access to financial resources, and increased resilience to market changes (Rodríguez-Gutiérrez et al., 2021; Garrido-Ruso et al., 2024; Murè et al., 2024). However, empirical evidence on the mechanisms that facilitate ESG adoption in SMEs remains fragmented. This highlights the need to identify the key drivers that can support SMEs in integrating sustainability principles into their strategic and operational activities.

2.2 Artificial Intelligence as a Driver of ESG Adoption

AI represents one of the most significant technological developments affecting modern organizations. By enabling companies to analyze large volumes of data, identify patterns, and automate decision-making processes, AI can improve operational efficiency, innovation capacity, and strategic responsiveness (Ancillai et al., 2023; Cheng et al., 2024). AI can support companies in monitoring environmental performance, optimizing resource utilization, and improving transparency in governance processes. Digital technologies also facilitate the collection and analysis of sustainability-related data, allowing companies to better assess their environmental and social impact and to respond more effectively to stakeholder expectations (Trueba-Castañeda et al., 2024; Tutore et al., 2024). AI can contribute to enhancing companies' ability to innovate and adapt their business models to sustainability challenges. By supporting predictive analytics and data-driven decision-making, AI enables companies to anticipate market trends and identify new opportunities related to sustainable products and processes. Despite these potential benefits, the adoption of AI may involve significant initial investments and implementation challenges, particularly for SMEs. Nevertheless, its role in improving decision-making processes and supporting sustainability-oriented strategies suggests that AI can be a key driver of ESG adoption and competitiveness.

2.3 Sustainability Management

SM refers to the integration of environmental and social considerations into corporate strategies, managerial practices, and organizational processes. It involves the development of policies, routines, and governance mechanisms aimed at aligning business activities with sustainability objectives and stakeholder expectations (Abad-Segura et al., 2020; Adewumi et al., 2024). SM encompasses a wide range of practices, including the implementation of environmental certifications, the adoption of sustainability reporting systems, the redesign of production processes to reduce environmental impact, and the development of training programs related to sustainability issues. Through these practices, companies can embed sustainability principles into their operational and strategic activities. Existing literature suggests that SM plays a crucial role in facilitating ESG adoption. Companies that develop structured sustainability-oriented processes are more likely to effectively implement ESG practices and align their strategies with broader societal and environmental goals (Lei and Yu, 2023). However, the relationship between SM and companies' competitive outcomes remains debated. While some studies highlight its positive impact on performance, others emphasize the potential organizational costs and challenges associated with sustainability implementation, particularly in resource-constrained contexts such as SMEs.

2.4 Competitive perspective in SMEs

In this study, competitive perspectives refer to a firm's perceived ability to maintain and strengthen its competitive position through innovation, adaptability, and long-term value creation. Rather than focusing exclusively on short-term financial performance, this concept captures broader dimensions of competitiveness, including market development, organizational resilience, and the capacity to innovate. Previous studies suggest that sustainability-oriented practices can contribute to improving firms' competitive positioning by enhancing reputation, strengthening stakeholder trust, and supporting innovation processes (Garrido-Ruso et al., 2024). However, the relationship between ESG practices and competitiveness is not always straightforward, especially in the context of SMEs. Resource limitations and structural constraints may influence the extent to which sustainability initiatives translate into tangible competitive advantages. Understanding how ESG practices affect SMEs' competitive perspectives therefore represents a relevant research challenge.

2.5 Research model and Hypotheses

Building on the existing literature, this study conceptualizes AI and SM as key drivers that support the adoption of ESG practices and influence companies' competitive perspectives.

AI can enhance companies' ability to process information, improve decision-making, and support sustainability-related initiatives. AI is expected to facilitate the implementation of ESG practices across their environmental, social, and governance dimensions. In addition, by improving operational efficiency and innovation capacity, AI may also directly contribute to companies' competitive perspectives. SM represents another important driver of ESG adoption. By embedding sustainability principles into organizational routines and managerial processes, companies are more likely to effectively implement ESG practices. However, its direct impact on competitive perspectives remains uncertain, as competitive benefits may emerge primarily through the effective implementation of sustainability practices. Finally, ESG practices themselves may influence companies' competitive perspectives. The environmental dimension may improve resource efficiency and reduce environmental risks, the social dimension may strengthen stakeholder relationships and employee engagement, and the governance dimension may enhance transparency and decision-making processes.

Based on these considerations, the following hypotheses are proposed:

- *H1.1 AI positively influences the environmental dimension of ESG.*
- *H1.2 AI positively influences the social dimension of ESG.*
- *H1.3 AI positively influences the governance dimension of ESG.*
- *H1.4 AI positively influences companies' competitive perspectives.*
- *H2.1 SM positively influences the environmental dimension of ESG.*
- *H2.2 SM positively influences the social dimension of ESG.*
- *H2.3 SM positively influences the governance dimension of ESG.*
- *H2.4 SM positively influences companies' competitive perspectives.*
- *H3.1 The environmental dimension of ESG positively influences companies' competitive perspectives.*
- *H3.2 The social dimension of ESG positively influences companies' competitive perspectives.*
- *H3.3 The governance dimension of ESG positively influences companies' competitive perspectives.*

We propose a conceptual model that integrates AI and SM as key drivers of ESG practices, and examine their effects on companies' competitive perspectives. The model captures both the influence of technological and managerial factors on ESG adoption and the role of ESG dimensions in shaping competitive outcomes.

Figure 1 shows the research model and the hypothesized relationships among the constructs. The model is tested empirically to assess the validity of the proposed relationships.

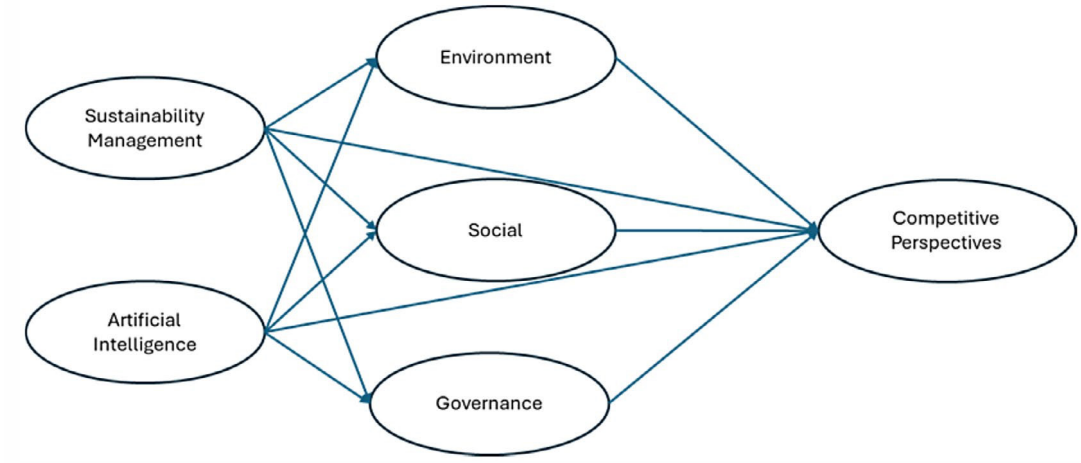


Figure 1: Conceptual model and hypothesized relationships.

3 Methodology

3.1 Sample and Data Collection

The empirical analysis is based on a sample of innovative SMEs operating in Italy. The firms were selected from the official national register of Innovative SMEs, established to support technologically advanced and research-oriented companies. These firms are characterized by systematic investments in research and development, the presence of intellectual property, and a high level of human capital, which are typical features of innovative SMEs. Such features make them a suitable context for analyzing the relationship between technological capabilities, sustainability practices, and competitive perspectives. To ensure greater consistency and reduce heterogeneity, the analysis focuses on SMEs located in the Lazio region. This region is considered representative of the national economic context in terms of industrial structure and development level. At the time of data collection, the regional register included 394 innovative SMEs. After excluding companies with incomplete contact information, the final population consisted of 355 companies. From this population, a random sample of 200 firms was selected. Data were collected through a structured online questionnaire administered to company representatives involved in strategic and innovation-related decision-making processes. The questionnaire consisted of closed-ended questions measured on a five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). The questionnaire was validated through expert evaluation involving both academics and SME managers. A pilot test was also conducted to ensure clarity and relevance of the items. The final dataset includes 163 valid responses, corresponding to a response rate of approximately 81%.

3.2 Measurement of Variables

The constructs included in the research model were operationalized using multiple indicators derived from established measures in the literature on ESG practices, digital innovation, and sustainability management (Abad-Segura et al., 2020; Cheng et al., 2024; Lei and Yu, 2023).

Each latent construct was measured through a set of manifest variables capturing different dimensions of firms' technological, managerial, and sustainability-related practices. AI was measured through indicators related to the adoption of digital technologies and structured innovation management practices, reflecting the companies' capability to support data-driven decision-making (Ancillai et al., 2023; Cheng et al., 2024). SM was operationalized through indicators capturing the presence of sustainability-oriented strategies, policies, and managerial practices within the organization (Adewumi et al., 2024).

The ESG construct was modelled as a multidimensional concept composed of three distinct dimensions: environmental, social, and governance practices, as commonly adopted in the literature (Jiang et al., 2024). The

environmental dimension includes actions related to emissions reduction, energy efficiency, resource management, renewable energy use, and circular economy initiatives (Trueba-Castañeda et al., 2024). The social dimension captures practices related to diversity and equality, employee training, workplace safety, inclusion, and employee involvement. The governance dimension includes practices related to sustainability communication, stakeholder engagement, certifications, and sustainable innovation (Cardinali and De Giovanni, 2022). Competitive perspectives refer to the companies' perceived ability to strengthen their competitive position through sustainable strategic orientation, market development, and organizational resilience (Garrido-Ruso et al., 2024; Murè et al., 2024). All items were measured using a five-point Likert scale.

3.3 Data Analysis

The research model was tested using Partial Least Squares Structural Equation Modelling (PLS-SEM), implemented through SmartPLS software. PLS-SEM was selected for several reasons. First, it is particularly suitable for analyzing complex models involving multiple latent constructs and relationships among variables. Second, this approach is appropriate for exploratory research contexts and does not require strict assumptions of multivariate normality. Third, PLS-SEM is well suited for studies with relatively small sample sizes (Hair et al., 2019). The analysis followed a two-step procedure commonly adopted in structural equation modelling (Hair et al., 2019). In the first step, the measurement model was assessed to evaluate the reliability and validity of the constructs through indicators such as Cronbach's alpha, composite reliability, and average variance extracted (AVE). In the second step, the structural model was analyzed to test the hypothesized relationships among the constructs. The explanatory power of the model was evaluated using the coefficient of determination (R^2), while the significance of the relationships was assessed through a bootstrapping procedure with 6000 resamples.

4 Results

4.1 Measurement Model Assessment

The evaluation of the empirical model followed the standard two-step approach adopted in PLS-SEM. The results indicate that all constructs meet the recommended thresholds, with values generally exceeding 0.7, thus confirming satisfactory reliability. SM shows a relatively low Cronbach's alpha (0.353), even if it still fails within the range deemed acceptable by some scholars (Bonett and Wright, 2015). However, composite reliability exceeds the recommended threshold, and the construct demonstrates acceptable convergent validity, suggesting that the measurement remains adequate for exploratory analysis (Hair et al., 2019). Indicator reliability was assessed by analyzing factor loadings. Most of the indicators show loadings above the acceptable threshold, indicating that they adequately represent their respective latent constructs. A limited number of indicators exhibit slightly lower values, which remain acceptable in exploratory research contexts. Convergent validity was evaluated through the Average Variance Extracted (AVE). The AVE values for all constructs exceed the recommended threshold of 0.50, confirming that the constructs explain a sufficient proportion of variance in their indicators. Overall, the results of the measurement model assessment support the adequacy of the constructs and allow the analysis to proceed to the evaluation of the structural model. Table 1 reports the results of the measurement model assessment. All constructs show satisfactory levels of internal consistency reliability, with Cronbach's alpha and composite reliability values exceeding the recommended threshold of 0.7. In addition, the AVE values are above 0.50 for all constructs, confirming convergent validity.

Table 1. The results of the measurement model assessment.

Construct	Cronbach's Alpha	Composite Reliability	AVE
Artificial Intelligence (AI)	0.729	0.881	0.787
Sustainability Management (SM)	0.353	0.756	0.607
Environmental Practices (E)	0.798	0.860	0.559
Social Practices (S)	0.874	0.904	0.612
Governance Practices (G)	0.801	0.863	0.561
Competitive Perspectives (CP)	0.742	0.852	0.657

4.2 Structural Model Assessment

After confirming the reliability and validity of the measurement model, the structural model was analyzed to test the hypothesized relationships among the constructs. The explanatory power of the model was assessed

using the coefficient of determination (R^2) for the endogenous variables. The results indicate a satisfactory level of explained variance, particularly for competitive perspectives, which shows a relatively high R^2 value. The significance of the relationships was evaluated through a bootstrapping procedure with 6000 examples. The findings show that AI has a positive and statistically significant effect on all three ESG dimensions supporting hypotheses H1.1, H1.2, and H1.3. AI also exhibits a significant direct effect on competitive perspectives, supporting hypothesis H1.4. SM shows positive and significant effects on the ESG dimensions, confirming hypotheses H2.1, H2.2, and H2.3. However, the direct relationship between SM and competitive perspectives is not statistically significant, leading to the rejection of hypothesis H2.4. Regarding the effects of ESG dimensions on competitive perspectives, the results reveal heterogeneous outcomes. The environmental and governance dimensions exert a positive and significant influence, supporting hypotheses H3.1 and H3.3. However, the environmental dimension shows a marginally significant effect on competitive perspectives. In contrast, the social dimension does not show a statistically significant effect, and therefore hypothesis H3.2 is not supported. Table 2 presents the estimated path coefficients, standard errors and p-value and a summary of the hypothesis testing results, indicating the direction and significance of each relationship included in the model.

Table 2. Relationships, estimated Path coefficients, standard deviation, p-value and summary of the hypothesis testing results.

Hypothesis	Relationship	Path coefficients	Standard deviation	T statistics	P-values	Confirmed/ Not confirmed
H1.1	AI → Environmental ESG	0.403	0.064	6.275	0.000	YES
H1.2	AI → Social ESG	0.320	0.074	4.327	0.000	YES
H1.3	AI → Governance ESG	0.487	0.060	8.137	0.000	YES
H1.4	AI → Competitive Perspectives	0.592	0.053	11.216	0.000	YES
H2.1	SM → Environmental ESG	0.365	0.072	5.080	0.000	YES
H2.2	SM → Social ESG	0.518	0.066	7.810	0.000	YES
H2.3	SM → Governance ESG	0.353	0.066	5.345	0.000	YES
H2.4	SM → Competitive Perspectives	-0.005	0.072	0.076	0.940	NO
H3.1	Environmental ESG → Competitive Perspectives	0.110	0.064	1.716	0.086	YES
H3.2	Social ESG → Competitive Perspectives	-0.079	0.086	0.925	0.355	NO
H3.3	Governance ESG → Competitive Perspectives	0.239	0.089	2.680	0.007	YES

5 Discussion

The aim of this study was to examine the role of AI and SM in fostering ESG practices and enhancing competitive perspectives among innovative SMEs. The results highlight that AI positively influences all three ESG dimensions confirming its relevance as a key enabler of sustainability-oriented decision-making. By supporting data processing, predictive analytics, and automation, AI allows companies to improve the monitoring of environmental performance, enhance transparency in governance processes, and support socially responsible practices. In addition, AI shows a direct positive effect on competitive perspectives. This finding suggests that AI contributes more broadly to companies' ability to innovate, adapt, and strengthen their competitive positioning. In the context of SMEs, AI can represent a critical lever for enhancing both sustainability and competitiveness. The results also confirm that SM contributes to the adoption of ESG practices across all three dimensions. Companies that integrate sustainability principles into their managerial processes and organizational routines are more likely to implement ESG practices effectively. This finding reinforces the importance of structured sustainability-oriented approaches in translating sustainability objectives into concrete actions. However, the analysis shows that SM does not exert a direct effect on competitive perspectives. This suggests that the mere presence of sustainability-oriented managerial frameworks is not sufficient to generate competitive advantages. Instead, competitive benefits seem to emerge when sustainability principles are effectively translated into ESG practices. In this sense, ESG adoption appears to play a mediating role in the relationship between SM and competitiveness. The findings reveal that the impact of ESG dimensions on competitive perspectives is not uniform. The environmental and governance dimensions show a positive and significant effect, whereas the social dimension does not exhibit a statistically significant influence. One possible explanation may be related to the characteristics of the analyzed sample, composed of innovative SMEs with relatively high levels of human capital and collaborative organizational practices. In such contexts, social initiatives may already be embedded in companies' routines and therefore may not represent a distinctive source of competitive advantage. Overall,

the results suggest that the relationship between sustainability and competitiveness in SMEs is shaped by the interaction between technological capabilities and managerial practices. AI enhances firms' ability to process information and support decision-making, while SM provides the organizational framework necessary to implement sustainability strategies. ESG practices represent the point where these elements converge, translating strategic intentions into concrete outcomes that can influence competitive positioning. The study highlights that ESG implementation should be considered as a strategic process that can support companies' long-term development.

6 Conclusions

In relation to the research questions, the findings provide clear answers to both inquiries. With regard to RQ1, the results show that AI and SM jointly support the adoption of ESG practices in SMEs, although through different mechanisms. While AI directly enhances companies' ability to implement ESG initiatives, SM contributes by embedding sustainability principles into organizational processes. Concerning RQ2, the results indicate that ESG adoption contributes to enhancing companies' competitive perspectives, although the effect varies across dimensions. In particular, environmental and governance practices emerge as significant drivers of competitiveness, whereas the social dimension does not show a direct impact in the analyzed context. The findings offer several implications for managers and policymakers. First, the results emphasize the strategic importance of AI in supporting sustainability-oriented innovation. Companies should consider AI not only as a technological investment but as a key enabler of data-driven decision-making and sustainability performance improvement. Second, managers should develop structured SM practices that integrate ESG principles into organizational processes, governance mechanisms, and strategic planning. However, the results suggest that such practices must be effectively translated into concrete ESG initiatives in order to generate tangible competitive benefits. Third, policymakers should promote initiatives aimed at facilitating SMEs' adoption of both digital technologies and sustainability practices. Financial incentives, training programs, and knowledge-sharing platforms may help SMEs overcome resource constraints and accelerate the implementation of ESG strategies. Regarding the limitations, the analysis focuses on innovative SMEs located in a specific geographical context. Future studies could extend the analysis to companies operating in different regions to compare sustainability adoption patterns across diverse contexts. The study does not explicitly include control variables such as firm size, industry characteristics, ownership structure, or degree of internationalization. Incorporating these factors in future research could provide a more comprehensive understanding of the determinants of ESG adoption and competitiveness. Future research could also explore additional drivers of ESG adoption, such as organizational culture, and external collaboration networks, to better understand how SMEs can effectively integrate sustainability into their business models.

7 Ethics declaration

The authors confirm that the research was conducted in accordance with ethical standards. Participation in the survey was voluntary and anonymous.

8 AI declaration

The authors used AI-assisted technologies to enhance the stylistic quality of the text via grammar correction, and language refinement. After using AI tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

References

Abad-Segura, E., González-Zamar, M.D. Infante Moro, J.C. and García, G. (2020.) Sustainable Management of Digital Transformation in Higher Education: Global Research Trends. *Sustainability* 12: 2107.

Adewumi, A. S., Opoku, A. and Dangana, Z. (2024). Sustainability Assessment Frameworks for Delivering Environmental, Social, and Governance (ESG) Targets: A Case of Building Research Establishment Environmental Assessment Method (BREEAM) UK New Construction. *Corporate Social Responsibility and Environmental Management* 31: 1– 3791.

Ancillai, C., Sabatini, A., Gatti, M. and Perna, A. (2023) Digital technology and business model innovation: A systematic literature review and future research agenda. *Technological Forecasting and Social Change*, 188, 122307.

Bonett, D.G. and Wright, T.A. (2015). Cronbach's Alpha Reliability: Interval Estimation, Hypothesis Testing, and Sample Size Planning. *Journal of Organizational Behaviour* 36: pp.3-15.

Cardinali, P.G. and De Giovanni, P. (2022) Responsible digitalization through digital technologies and green practices. *Corporate Social Responsibility and Environmental Management*, 29(4), pp. 984–995.

Cheng, J., Xu, N. R., Khan, N. U. and Singh, H. S. M. (2024). The Impacts of Artificial Intelligence Literacy, Green Absorptive Capacity, and Green Information System on Green Innovation. *Corporate Social Responsibility and Environmental Management* 32: pp. 2375–2389.

Garrido-Ruso, M., Otero-González, L., López-Penabad, M.-C. and Durán Santomil, P. (2024). Does ESG Implementation Influence Performance and Risk in SMEs? *Corporate Social Responsibility and Environmental Management* 31, no. 5: pp. 4227–4247.

Hair, J.F., Hult, G.T.M., Ringle, C.M. and Sarstedt, M. (2019) When to use and how to report the results of PLS-SEM. *European Business Review*, 31(1), pp. 2–24.

Jiang, P.-C., Feng, G.-F., Wang, H.-J. and Chang, C.-P. (2024.) CSR From Different Perspectives: The Global ESG Indexes Updated. *Corporate Social Responsibility and Environmental Management* 31: pp. 4694–4714.

Lei, X. and Yu, J. (2023) Striving for sustainable development and corporate ESG performance. *Corporate Social Responsibility and Environmental Management*, 31, pp. 1177–1202.

Murè, P., Giorgio, S., Antonelli, V. and Crisafulli, A. (2024) ESG assessment and competitiveness of European SMEs. *Frontiers in Environmental Economics*, 3, 1452416.

Rodríguez-Gutiérrez, P., Guerrero-Baena, M.D., Luque-Vílchez, M. and Castilla-Polo, F. (2021) Sustainability reporting decision-making in SMEs. *Journal of Environmental Planning and Management*, 64(14), pp. 2618–2640.

Rubino, M., Mastrorocco, I. and Garegnani, G.M. (2024). The Influence of Market and Institutional Factors on ESG Rating Disagreement. *Corporate Social Responsibility and Environmental Management*, 31, no. 5: pp. 3916–3926.

Trueba-Castañeda, L., Somohano-Rodríguez, F.M. and Torre-Olmo, B. (2024). Does Digitalisation Enable Small and Medium-Sized Enterprises to Become More Sustainable? *Corporate Social Responsibility and Environmental Management* 31, no. 5: pp. 4893–4909.

Tutore, I., Parmentola, A., Costagliola di Fiore, M. and Calza, F. (2024) Artificial intelligence effects on circular economy actions. *Corporate Social Responsibility and Environmental Management*, 31(5), pp. 4772–4782.

AI-Enabled Knowledge Ecosystems: Conditions for Learning and Deskilling Across Economic Sectors

Maria Cristina Manocchio[1], Francesca Di Virgilio[1], Gianluigi Mangia[2], Antonia Puccio[1]

[1] Department of Economics, University of Molise, Campobasso, Italy

[2] Department of Economics, Management, Institutions, University of Naples Federico II, Naples, Italy

Abstract

In modern organizations, the integration of artificial intelligence is revolutionizing how knowledge is created, shared, and utilized, while also altering the dynamics of organizational learning, skill enhancement, and governance frameworks. Generative AI is pivotal in making explicit knowledge more accessible and in optimizing both automation and decision-making processes. Nevertheless, it introduces challenges, such as the risk of diminishing tacit knowledge and an increasing reliance on algorithm-driven insights. Thus, AI is gradually becoming both a technological foundation and an active player in organizational knowledge environments. The effectiveness of AI in enhancing knowledge creation and innovation, or conversely, in causing skill mismatches and deskilling, is largely influenced by organizational learning capabilities, culture, and governance mechanisms. Although there is increasing academic interest, there remains a lack of empirical evidence on the influence of AI on organizational knowledge ecosystems beyond individual companies. This study addresses this gap by examining how various AI adoption configurations affect organizational knowledge ecosystems and learning dynamics at the sectoral level. The research employs secondary data from the 2025 Survey on Information and Communication Technologies in Enterprises by ISTAT, concentrating on Italian companies with at least 10 employees. Data at the firm level is aggregated by economic sector, allowing for a meso-level analysis of AI adoption patterns. From a capability-based viewpoint, the study distinguishes between AI technologies that are cognitively oriented, aiding in knowledge creation, and those that are automation-oriented, primarily focused on task substitution. A sector-level cluster analysis is conducted to pinpoint distinct AI adoption patterns and their integration within organizations. This research aims to reveal various organizational knowledge ecosystems, each defined by a unique balance between AI applications focused on learning and those centered on automation. By employing a configurational and sector-specific approach, the study enhances the knowledge management literature by clarifying the circumstances under which AI adoption promotes organizational learning and the development of sustainable capabilities, rather than leading to a reduction in skills.

Keywords: Artificial Intelligence, Knowledge Ecosystems, Organizational Learning, Deskilling, Knowledge Management.

Paper type: Academic Research Paper

1 Introduction

Across contemporary organizations, the adoption of artificial intelligence (AI) is reshaping how knowledge is created, shared, and applied, while simultaneously reconfiguring learning dynamics, skills, and governance structures (Alavi, Leidner and Mousavi, 2024; Leoni et al., 2024). Recent investigations indicate that AI, particularly generative AI (GenAI), fundamentally transforms knowledge management processes by enabling rapid access to explicit knowledge and supporting cognitive tasks such as analysis, forecasting, and decision-making (Lin and Wu, 2025; Dey et al., 2024; Leoni et al., 2024; Mikalef et al., 2023). Concurrently, these advancements elicit concerns regarding the potential erosion of tacit, experience-based knowledge and an excessive dependence on algorithmic outputs (Storey, 2025; Alavi, Leidner and Mousavi, 2024). AI adoption can generate both opportunities and tensions: it can enhance knowledge creation, learning, and innovation performance through human–AI collaboration, yet automation-oriented applications may induce deskilling, skill mismatches, and the need for reskilling (Arnaout et al., 2025; Cheng et al., 2025; Rostamzadeh et al., 2025; Cramarenco, Burcă-Voicu and Dabija, 2023). Organizational learning capabilities moderate these effects by converting individual gains into team-level advantages and dynamic capabilities (Han et al., 2025; Jamil et al., 2025; Yan, Husted and Fath, 2025). Proper governance, leadership, and skill development are critical for AI to support valuable learning activities and sustain competitive advantage (Schwaeke et al., 2025; Storey, 2025; Dey et al., 2024).

Despite these advancements, empirical evidence regarding the influence of AI adoption on organizational knowledge ecosystems beyond the firm level remains scarce (Dahlke et al., 2024; Leoni et al., 2024). Existing research has predominantly concentrated on firm-level adoption or performance outcomes, neglecting the

sectoral patterns through which AI diffusion impacts learning dynamics and knowledge management practices. This study expands upon this gap by examining the impact of varying AI adoption configurations on the learning dynamics and organizational knowledge ecosystems at the sectoral level. The research specifically addresses the following question:

- *RQ1: Under what conditions does AI adoption support organizational learning rather than deskilling within organizational knowledge ecosystems?*

This research utilizes the knowledge-based perspective of the firm alongside organizational learning theory to examine how diverse AI adoption strategies shape distinct organizational knowledge ecosystems. By investigating the sectoral trends of AI adoption within the Italian industrial landscape, the research aims to shed light on the interaction between AI technologies, organizational practices, and data capabilities, and how these elements collectively influence knowledge creation and learning outcomes at the meso level.

2 Literature Review

2.1 AI Adoption and Organizational Knowledge Ecosystems

The deployment of AI is progressively incorporated into the comprehensive knowledge structures of enterprises, affecting how companies create, share, and apply expertise (Schwaeke et al., 2025). AI transcends being a mere technological enhancement; it reconfigures organizational knowledge ecosystems, defined as the interplay among digital infrastructures, human capabilities, governance mechanisms, and inter-organizational networks (Cheng et al., 2025; Zhang, Zuo and Yang, 2025). The incorporation of AI within the Technology–Organization–Environment (TOE) framework is influenced by technological attributes such as relative advantage and compatibility, as well as organizational factors including readiness, leadership, skills, and learning orientation (Brennecke et al., 2025; Cimino et al., 2025). Dynamic capabilities, such as adaptability, absorptive capacity, and knowledge integration, are often linked to increased AI adoption intensity and more sophisticated implementation strategies. Conversely, challenges such as financial constraints, intellectual property issues, low trust levels, regulatory ambiguities, and resistance to change impede adoption and may restrict AI to narrower, efficiency-driven applications (Badghish and Soomro, 2024; Booyse and Scheepers, 2024). Existing literature indicates that AI facilitates various organizational pathways. In particular contexts, AI augments knowledge management systems, converting them into adaptive frameworks that support real-time analytics, collaborative learning, and innovation (Abdelrahman et al., 2025). Conversely, in different contexts, AI is utilized mostly for process automation and task replacement, exhibiting limited integration into comprehensive learning frameworks (Bankins et al., 2024). The distinct pathways of AI adoption underscore that the generation of knowledge-intensive outcomes is not inherently guaranteed but is contingent upon the surrounding ecosystem, which includes skills, culture, and governance. Furthermore, knowledge ecosystems are dependent on social capital and inter-organizational networks, which facilitate the flow of tacit knowledge and enhance collective resilience (Kucharska and Erickson, 2023). Leadership, culture, and skills development impact the extent to which AI enhances collaborative knowledge dynamics or perpetuates fragmented, efficiency-oriented structures (Dey et al., 2024).

2.2 Dynamics of Organizational Learning and Deskilling: Conditions and Mechanisms

Organizational learning (OL) is defined as the evolving process by which an organization's knowledge, actions, and collective beliefs develop through experience and become ingrained in its routines, systems, and culture (Nawaz et al., 2025; Ademi, Sætre and Klungseth, 2024). This process involves the development, preservation, distribution, and application of knowledge across individual, team, and organizational levels. It can occur through gradual modifications (single-loop learning) or through more substantial shifts that challenge fundamental beliefs and strategies (double-loop learning) (Ademi, Sætre and Klungseth, 2024). Conversely, deskilling involves the reduction or inadequate use of human abilities due to technology assuming cognitive responsibilities. In AI-enhanced workplaces, deskilling may happen when automation replaces human decision-making, leading to a decline in critical thinking, experiential learning, and cognitive engagement, particularly among employees with less experience (Choudhury and Chaudhry, 2024). The integration of AI in digital transformation brings both the promise of enhanced learning and the threat of skill erosion. From an organizational learning standpoint, AI systems are increasingly becoming partners in the creation, retention, and distribution of knowledge, thereby supporting continuous cycles of knowledge acquisition and organizational memory at micro, meso, and macro levels (Gibson et al., 2023; Jarrahi et al., 2022). When embedded within learning infrastructures, AI strengthens dynamic capabilities by enabling organizations to identify, leverage, and

adapt resources in response to shifting environments (Han et al., 2025). The outcomes largely depend on whether AI substitutes or augments human cognition. Substitution may generate reduced expertise and passive monitoring, whereas augmentation-based human–AI collaboration sustains reflection and competence development (Bauer et al., 2025; Mosqueira-Rey et al., 2022). Consequently, OL acts as a mediating mechanism through which AI adoption supports sustainable performance, provided organizations invest in human capital, AI literacy, and active engagement with AI outputs, thereby preventing deskilling and fostering adaptive capabilities (Rostamzadeh et al., 2025; Alavi, Leidner and Mousavi, 2024).

2.3 Role of Organizational Culture, Skills and Governance in Sustainable AI Integration

The influence of organizational culture, skills, and governance is pivotal in determining whether AI integration boosts learning and capabilities or results in reliance and diminished skills. Examining AI from a governance standpoint involves perceiving it as a framework of rules, practices, and processes that must align with an organization's strategies, objectives, and values to be both effective and legitimate (Birkstedt et al., 2023). A culture that prioritizes learning, fostering an environment of experimentation, knowledge exchange, and psychological safety, is essential for enhancing human–AI collaboration and reducing resistance to AI integration. Trust and ethical awareness play a crucial role in determining employees' readiness to rely on algorithmic outputs while maintaining their critical judgment (Schwaeke et al., 2025; Bankins et al., 2024). The extent to which workers embrace AI tools is significantly influenced by their skills and human capital, particularly in areas such as reskilling, upskilling, data literacy, and AI literacy. Research indicates that AI provides the most value when algorithmic skills are widely distributed among domain experts, allowing them to interpret, challenge, and incorporate model outputs into decision-making processes (Tambe, 2025; Li, 2022). Failing to continuously invest in skill development can lead to greater skill mismatches and jeopardize well-being, as it increases dependence on complex systems (Cramarenco, Burcă-Voicu, and Dabija, 2023). The third pillar of integration is governance and leadership. By employing AI governance frameworks, oversight units, and establishing clear accountability, strategic alignment, change management, and ethical oversight transform theoretical AI principles into practical, enforceable actions and consistent routines (Nair et al., 2024; Birkstedt et al., 2023; Kelley, 2022).

3 Research Methodology

This study examines the implementation and sectoral trends of AI in Italian companies with a minimum of 10 employees, emphasizing the coexistence of various AI applications and organizational digital competencies across sectors. Considering the diverse use of AI in business operations, a quantitative exploratory methodology is employed to delineate sectoral trends and discern unique knowledge ecosystems, rather than to evaluate causal hypotheses. Three composite indexes were developed to encapsulate the intricacies of AI deployment. Cognitive AI includes technologies that enhance knowledge, decision-making, and innovation; Automation AI refers to applications intended for task replacement and workflow automation; and Embeddedness denotes organizational digital competencies, encompassing the use of ERP, CRM, and BI systems, data analytics, and overall digitalization levels. These indices provide descriptive statistics and cluster analysis, facilitating a systematic evaluation of sectoral differences in AI utilization and digital competencies.

3.1 Data Collection

This study is based on data from ISTAT (2025), drawn from the ICT–Enterprises tables. It looks at 26,246 Italian companies with at least 10 employees, analyzing them at the sector level, which is the main focus of the study.

3.2 Descriptive Statistics

Three theoretically based indexes were developed to offer a comprehensive perspective of AI adoption and digital capabilities across sectors: Cognitive AI, Automation AI, and Embeddedness. Cognitive AI includes AI technologies that enhance knowledge, decision-making, and innovation, such as text mining, natural language processing, machine learning, workflow automation, and research and development activities. Automation AI reflects applications aimed at task substitution and process automation, such as image and video generation, object recognition, and autonomous robotics. Embeddedness represents organizational digital capabilities, combining ERP/CRM/BI usage, data analytics activities, and overall sectoral digitalization.

Table 1 provides the complete identification of the variables employed in the formulation of each index, along with their corresponding source tables and descriptions. This thorough mapping guarantees openness in

index formulation and facilitates the repetition of the analysis. All variables are represented as sectoral percentages of enterprises implementing each technology. To ensure comparability across dimensions, variables were standardized using min-max scaling prior to aggregation. Each index was calculated as the unweighted arithmetic mean of its normalized constituents. Robustness assessments employing different weighting methodologies produced consistent outcomes.

Table 1: Variables used to construct AI and digital capabilities indice

Index	Source Table	Description
Cognitive AI	9A	Extract knowledge and information from text documents (text mining)
	9A	Convert spoken language into machine-readable format (speech recognition)
	9A	Generate written or spoken language or programming code (natural language generation)
	9A	Analyze data through machine learning, deep learning, and neural networks
	9A	Automate workflows or support decision-making (RPA / decision support)
	9B	Research & Development (R&D) or innovation activities
Automation AI	9A	Generate images, video, or audio content
	9A	Identify objects or people based on images (image recognition)
	9A	Enable autonomous movement of machines based on environment observation (robots, drones, vehicles)
Embeddedness	6	Use of ERP software
	6	Use of CRM software
	6	Use of Business Intelligence software
	7	Data analytics via internal employees or other organizations
	7	Data analytics via external organizations
	3	Very low/ Low/ High/ Very high level of digitalization

Summary statistics of the three indices at the sectoral level are reported in **Table 2,** showing that Cognitive AI exhibits considerable variation (median = 34.53%, max = 57.41%), suggesting strong use of knowledge-enhancing AI in certain sectors. Automation AI exhibits reduced values (median = 21.48%), suggesting that task substitution is less common, whereas Embeddedness demonstrates moderate levels (median = 28.75%), indicating a broad degree of organizational digital capabilities.

Table 2: Summary statistics of AI and digital capability indices

Index	Min	1st Qu.	Median	Mean	3rd Qu.	Max
Cognitive AI	25.32	30.22	34.53	36.12	41.65	57.41
Automation AI	11.78	17.78	21.48	22.70	26.37	40.32
Embeddedness	19.69	25.87	28.57	29.17	32.12	37.66

Figure 1 shows the distribution of the three indices by implementing boxplots, highlighting disparities among macro-sector groups and the presence of outliers.

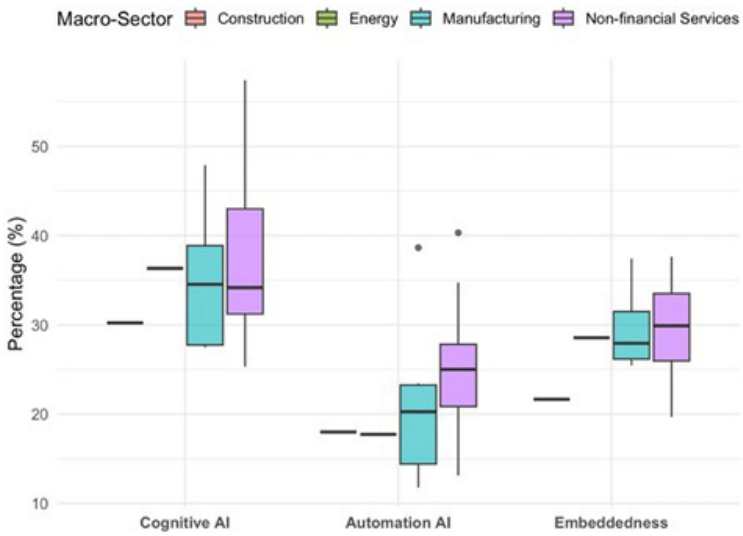


Figure 1: Distribution of AI and Digitalization Indices Across Sectors

Cognitive AI exhibits highest variability, with numerous sectors attaining levels over 55%, signifying robust utilization of AI for knowledge generation, decision-making assistance, and creativity. The prevalence of Automation AI is rather low, with median values about between 21% and 22%, indicating that the substitution of human tasks is not extensively adopted across many sectors. Embeddedness, which consolidates digitalization levels, ERP/CRM/BI utilization, and data analytics endeavors, exhibits a more concentrated distribution (median ~28.8%), suggesting that the majority of sectors possess a moderate level of organizational digital capabilities. The macro-sector color coding offers a prompt visual indication of which industries are at the forefront of AI adoption and which are underdeveloped, highlighting sectoral diversity and validating further cluster analysis to discern unique knowledge ecosystems.

3.3 Cluster Analysis

A cluster analysis was performed to identify patterns and classify sectors into distinct knowledge ecosystems. This analysis was based on the three indices of Cognitive AI, Automation AI, and Embeddedness, which are a widely used exploratory technique for grouping observations according to similarity across multiple dimensions (Hair et al., 2010). The cluster analysis revealed the fundamental patterns of AI deployment across several Italian sectors. The optimal number of clusters was determined using hierarchical clustering and silhouette analysis, resulting in a three-cluster solution that provided the clearest separation between sectoral knowledge ecosystems. Table 3 defines the sectoral distribution throughout the three clusters, specifying the composition of each cluster for macro-sector representation.

Table 3: Distribution of sectors across clusters by macro-sector

Cluster	Manufacturing	Energy	Construction	Non-Financial Services	Total
1	5	0	1	1	**7**
2	3	1	0	12	**16**
3	1	0	0	1	**2**

Cluster 1 is dominated by manufacturing sectors, with some representation from construction and non-financial services. Cluster 2 primarily comprises non-financial services, with a few manufacturing sectors included. Cluster 3 is small, including only one manufacturing and one service sector, highlighting the heterogeneous composition of sectoral knowledge ecosystems. Although Cluster 3 includes a limited number of sectors, silhouette scores confirmed the stability and internal coherence of the three-cluster solution.

A deeper examination of cluster profiles, reported in Table 4, shows differences in the mean, standard deviation, and median values of Cognitive AI, Automation AI, and Embeddedness across clusters.

Table 4: Cluster profiles: means, SD, and medians for Cognitive AI, Automation AI, and Embeddedness

Cluster	Mean Cognitive AI	SD Cognitive AI	Median Cognitive AI	Mean Automation AI	SD Automation AI	Mean Embeddedness	SD Embeddedness
1	47.60	5.95	46.80	26.50	6.05	33.50	3.70
2	31.80	4.37	31.80	19.20	5.34	28.30	3.68
3	30.40	6.95	30.40	37.50	3.95	20.90	1.69

Cognitive AI reaches its highest values in Cluster 1 (mean = 47.6%), indicating sectors where AI is predominantly leveraged for knowledge-intensive activities and innovation-oriented processes. Clusters 2 and 3 exhibit notably diminished values, suggesting limited cognitive applications. Automation AI exhibits a distinct pattern: Cluster 3 exhibits the highest mean (37.5%), indicating areas where AI deployment is predominantly focused on task automation. Clusters 1 and 2 have more moderate values, with Cluster 2 displaying the lowest intensity of automation-focused applications. Embeddedness displays a comparable gradient, with Cluster 1 demonstrating the highest level of digital integration (mean = 33.5%), Cluster 2 positioned at an intermediate level, and Cluster 3 indicating relatively lesser organizational digital capabilities.

On this basis, the clusters can be conceptually characterized as follows. Cluster 1, identified as "Cognitive Learning Ecosystems," is characterized by a high degree of cognitive intensity alongside significant digital integration. This cluster represents sectors where AI is extensively incorporated into processes of knowledge creation, decision support, and innovation. Cluster 2, identified as "Automation-Oriented Ecosystems," occupies an intermediate position, indicative of a more gradual and efficiency-focused adoption strategy that balances cognitive and automation-oriented applications. Cluster 3, referred to as "Emerging AI Ecosystems," is smaller and structurally distinct, marked by strong automation intensity but limited cognitive applications and weaker

digital integration. This suggests a concentrated focus on task substitution rather than comprehensive knowledge development.

Figures 2 and 3 further illustrate these patterns. Figure 2 shows the average values of the three indices for each cluster, emphasizing the disparities in AI adoption approaches.

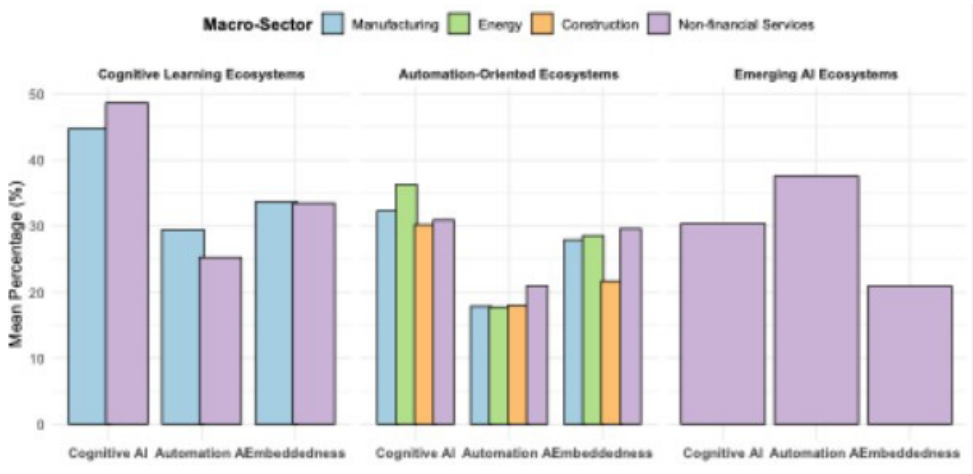


Figure 2: Mean values of Cognitive AI, Automation AI, and Embeddedness for each cluster with breakdown by macro-sector

The figure clearly differentiates the clusters and complements the numerical profiles reported in Table 3, reinforcing the sectoral heterogeneity and the existence of distinct knowledge ecosystems.

Figure 3 maps sectors along the Cognitive AI and Automation AI dimensions, revealing a clear spatial differentiation among clusters and their macro-sectoral composition.

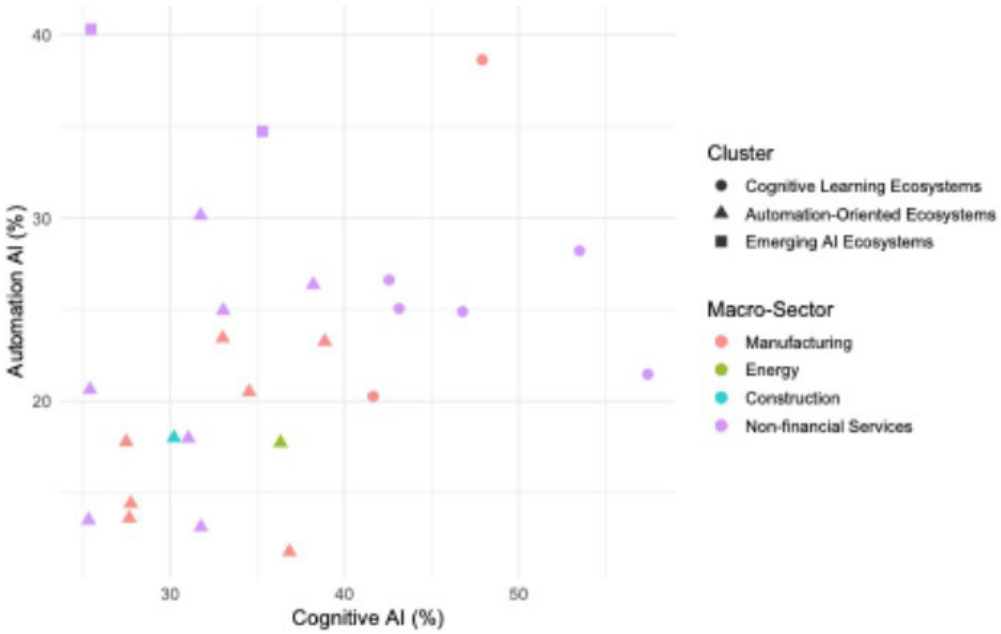


Figure 3: Sectoral Distribution of AI Adoption by Cluster and Macro-Sector

The scatterplot offers a clear spatial differentiation among clusters, so validating the internal coherence of the clustering solution. Cognitive Learning Ecosystems occupy the upper-left quadrant of the figure, distinguished by high levels of Cognitive AI and moderate levels of Automation AI, primarily linked to the Manufacturing and Non-financial Services sectors. Automation-Oriented Ecosystems hold a pivotal role, demonstrating intermediate values across all dimensions and a more diverse macro-sectoral composition. In contrast, Emerging AI Ecosystems are situated in the upper-right quadrant, characterized by high levels of Automation AI and little Cognitive AI, signifying a predominant emphasis on task automation over knowledge-

intensive applications. The figure illustrates the correlation between various AI adoption strategies and certain sectoral configurations, affirming the perspective of clusters as heterogeneous knowledge ecosystems.

3.1.1 Cluster Validation

A one-way analysis of variance (ANOVA) was performed for each of the three indices—Cognitive AI, Automation AI, and Embeddedness—to ascertain whether the identified clusters represent statistically distinct configurations. Table 5 illustrates statistically significant differences across clusters for all measured indices (Cognitive AI: $F = 25.70$, $p < 0.001$; Automation AI: $F = 12.18$, $p < 0.001$; Embeddedness: $F = 10.77$, $p < 0.001$). These results substantiate that the clusters are indicative of meaningful variations in AI adoption levels and organizational digital capabilities, rather than being arbitrary groupings.

Table 5: One-way ANOVA Results by Cluster

Index	F-statistic	p-value
Cognitive AI	25.70	< 0.001
Automation AI	12.18	< 0.001
Embeddedness	10.77	< 0.001

To investigate more thoroughly whether cluster pairs contributed to these differences, post-hoc multiple comparison tests utilizing Tukey's HSD were used (Tukey, 1949). The results are presented in Table 6, which demonstrates that there are substantial pairwise differences in Cognitive AI between Cluster 1 ("Cognitive Learning Ecosystems") and both Clusters 2 and 3. Automation AI also exhibits significant differences across all clusters, with Cluster 3 ("Emerging AI Ecosystems") exhibiting the highest values. Additionally, there are substantial disparities in embeddedness among all cluster pairings, which underscores a distinct division in the digital capabilities of organizations.

Table 6: Tukey HSD Post-Hoc Comparisons by Cluster

Index	Comparison	Mean difference	p-value
Cognitive AI	Cluster 2 – Cluster 1	−15.74	< 0.001
Cognitive AI	Cluster 3 – Cluster 1	−17.19	< 0.001
Cognitive AI	Cluster 3 – Cluster 2	−1.45	n.s.
Automation AI	Cluster 2 – Cluster 1	−7.25	0.021
Automation AI	Cluster 3 – Cluster 1	11.08	0.050
Automation AI	Cluster 3 – Cluster 2	18.33	< 0.001
Embeddedness	Cluster 2 – Cluster 1	−5.22	0.011
Embeddedness	Cluster 3 – Cluster 1	−12.64	< 0.001
Embeddedness	Cluster 3 – Cluster 2	−7.41	0.031

The aggregate evidence from ANOVA and post-hoc tests confirms the legitimacy of the cluster solution, indicating that the three clusters represent distinct sectoral knowledge ecosystems distinguished by varied adoption of Cognitive and Automation AI and various levels of Embeddedness.

4 Discussion

The examination of cluster characteristics and sectoral distributions highlights diverse configurations of AI adoption within Italian organizations, underscoring the importance of identifying distinct knowledge ecosystems. The clusters exhibit a non-random distribution, revealing unique macro-sectoral patterns that underscore the variations in AI adoption strategies and digital competencies. The Cognitive Learning Ecosystems cluster is primarily associated with knowledge-intensive manufacturing activities and advanced non-financial services. In these sectors, AI is embedded in processes of knowledge creation, decision support, and innovation. Increased levels of Cognitive AI, along with strong digital integration, signify organizational environments that promote learning, improve the collaboration between human talents and AI technologies, and support the advancement of technology capabilities. The Automation-Oriented Ecosystems cluster displays a broader and more heterogeneous sectoral composition, including traditional manufacturing, construction, and several non-financial services. The integration of AI within these sectors is both progressive and equitable, reflecting a systematic incorporation of AI into established production and organizational practices. This trend suggests a practical utilization of AI, largely focused on improving efficiency and partially replacing tasks, rather than completely revolutionizing knowledge-intensive processes. The Emerging AI Ecosystems cluster is relatively small and concentrated within a limited number of sectors. It is characterized by a high degree of Automation

AI, yet exhibits comparatively low levels of Cognitive AI and Embeddedness. These sectors primarily focus on task automation with limited integration into broader digital and organizational infrastructures. Although this strategy may result in short-term productivity gains, it presents fewer opportunities for learning and skill enhancement, thereby raising potential concerns regarding deskilling dynamics.

These findings emphasize the variability of AI adoption across business sectors and highlight the necessity of considering both technological and organizational skills when assessing the potential for knowledge creation and skill development. The identification of distinct industry knowledge ecosystems offers an analytical framework for comprehending diverse AI trajectories and their effects on skills and organizational learning. These findings suggest that AI adoption does not follow a uniform technological trajectory but is embedded within sector-specific organizational and knowledge structures that shape its developmental path.

5 Conclusion

The rapid diffusion of AI is increasingly reshaping how organizations generate, integrate, and exploit knowledge across economic sectors (Alavi, Leidner and Mousavi, 2024; Leoni et al., 2024; Storey, 2025). Within this evolving landscape, understanding how different configurations of AI adoption influence organizational learning and knowledge dynamics has become a central issue in management and knowledge management research. Drawing on sectoral data from Italian enterprises, this study identified three distinct configurations of AI adoption—Cognitive Learning Ecosystems, Automation-Oriented Ecosystems, and Emerging AI Ecosystems—based on the interaction between Cognitive AI, Automation AI, and organizational digital embeddedness. The empirical results highlight that AI adoption does not follow a uniform trajectory across sectors. Rather, different technological and organizational configurations shape alternative knowledge ecosystems, influencing whether AI supports knowledge creation and organizational learning or primarily drives efficiency-oriented automation. These findings reinforce prior research emphasizing that the outcomes of AI adoption depend not only on technological capabilities but also on organizational structures, skills, and governance mechanisms (Birkstedt et al., 2023; Bankins et al., 2024). Overall, the analysis suggests that sectors characterized by stronger digital infrastructures and cognitively oriented AI applications are more likely to foster learning-oriented ecosystems, while automation-focused configurations may generate productivity gains but potentially limit opportunities for knowledge development. Future research could extend this analysis by examining firm-level adoption patterns and exploring cross-country comparisons, in order to better understand how institutional contexts, organizational capabilities, and governance practices shape the long-term relationship between AI adoption, organizational learning, and knowledge ecosystems.

6 Ethical considerations

The protection of the privacy of research participants has been ensured. Anonymity of individuals and organisations participating in the research has been ensured. There are no possible conflicts of interests. All the previously published work along with their authors have been duly acknowledged in references. There is no use of offensive, discriminatory, or other unacceptable language in the paper.

7 AI declaration

The authors have not used an AI tool in developing the paper.

References

Abdelrahman, M., Papamichail, N., Hammady, R., Kurt, Y., Abdallah, W., Liu, L., and Kapser, S. (2025) "Organisational Culture and Knowledge Management Systems Adoption: Examining Employee Usage in Multinational Corporations," *Global Journal of Flexible Systems Management*, 26(3), pp. 439–467.

Ademi, B., Sætre, A. and Klungseth, N. (2024) "Advancing the understanding of sustainable business models through organizational learning", *Business Strategy and the Environment*, Vol. 33, No. 6, pp. 5174–5194.

Alavi, M., Leidner, D. and Mousavi, R. (2024) "Knowledge Management Perspective of Generative Artificial Intelligence", *Journal of the Association for Information Systems*, Vol. 25, No. 1, pp. 1–12.

Arnaout, B., El Nemar, S., Kokkinopoulou, E., Skaf, Y., Vrontis, D., Esposito, M. and Rebeiz, K.S. (2025) "Bridging AI and Human Intelligence: Advancing Knowledge Ecosystems Through Effective Communication for Strategic Innovation in the US Utility Industry", *Journal of Knowledge Management*, pp. 1–21.

Badghish, S. and Soomro, Y.A. (2024) "Artificial Intelligence Adoption by SMEs to Achieve Sustainable Business Performance: Application of Technology–Organization–Environment Framework," *Sustainability*, 16(5), p. 1864.

Bankins, S., Ocampo, A., Marrone, M., Restubog, S., and Woo, S. (2024) "A multilevel review of artificial intelligence in organizations: Implications for organizational behavior research and practice," *Journal of Organizational Behavior*, 45(2), pp. 159–182.

Bauer, E., Greiff, S., Graesser, A., Scheiter, K. and Sailer, M. (2025) "Looking Beyond the Hype: Understanding the Effects of AI on Learning", *Educational Psychology Review*, Vol. 37.

Birkstedt, T., Minkkinen, M., Tandon, A. and Mäntymäki, M. (2023) "AI Governance: Themes, Knowledge Gaps and Future Agendas", *Internet Research*, Vol. 33, No. 7, pp. 133–167.

Booyse, D. and Scheepers, C.B. (2024) "Barriers to adopting automated organisational decision-making through the use of artificial intelligence," *Management Research Review*, 47(1), pp. 64–85.

Brennecke, J., Coutinho, J., Gilding, M., Lusher, D., and Schaffer, G. (2025) "Invisible Iterations: How Formal and Informal Organization Shape Knowledge Networks for Coordination," *Journal of Management Studies*, 62(2), pp. 706–747.

Cheng, Q., Zhang, S., Wang, L., and Lu, Y. (2025) "Inter-organizational knowledge sharing in the age of artificial intelligence: the mediating role of knowledge digitization and cross-language collaboration", *Journal of Knowledge Management*, *29*(8), 2724-2744.

Choudhury, A. and Chaudhry, Z. (2024) "Large Language Models and User Trust: Consequence of Self-Referential Learning Loop and the Deskilling of Health Care Professionals", *Journal of Medical Internet Research*, Vol. 26.

Cimino, A., Corvello, V., Troise, C., Thomas, A., and Tani, M. (2025) "Artificial Intelligence Adoption for Sustainable Growth in SMEs : An Extended Dynamic Capability Framework," *Corporate Social Responsibility and Environmental Management*, 32(5), pp. 6120–6138

Cramarenco, R.E., Burcă-Voicu, M.I. and Dabija, D.C. (2023) "The Impact of AI(AI) on Employees' Skills and Well-Being in Global Labor Markets: A Systematic Review", *Oeconomia Copernicana*, Vol. 14, No. 3, pp. 731–767.

Dahlke, J., Beck, M., Kinne, J., Lenz, D., Dehghan, R., Wörter, M. and Ebersberger, B. (2024) "Epidemic effects in the diffusion of emerging digital technologies: evidence from AIadoption", *Research Policy*, Vol. 53, No. 2, Article 104917.

Dey, P.K., Chowdhury, S., Abadie, A., Yaroson, E., and Sarkar, S. (2024) "Artificial intelligence-driven supply chain resilience in Vietnamese manufacturing small- and medium-sized enterprises," *International Journal of Production Research*, 62(15), pp. 5417–5456.

Gibson, D., Kovanović, V., Ifenthaler, D., Dexter, S. and Feng, S. (2023) "Learning theories for AI promoting learning processes", *British Journal of Educational Technology*, Vol. 54, No. 5, pp. 1125–1146.

Hair Jr., J.F., Black, W.C., Babin, B.J. and Anderson, R.E. (2010) *Multivariate Data Analysis*, 7th ed., Upper Saddle River, NJ: Pearson Prentice Hall.

Han, S., Zhang, D., Zhang, H. and Lin, S. (2025) "AITechnology, Organizational Learning Capability, and Corporate Innovation Performance: Evidence from Chinese Specialized, Refined, Unique, and Innovative Enterprises", *Sustainability*, Vol. 17, No. 6.

Jamil, K., Zhang, W., Anwar, A. and Mustafa, S. (2025) "Exploring the Influence of AI Adoption and Technological Readiness on Sustainable Performance in Pakistani Export Sector Manufacturing Small and Medium-Sized Enterprises", *Sustainability*, Vol. 17, No. 8, 3599.

Jarrahi, M., Askay, D., Eshraghi, A. and Smith, P. (2022) "AI and knowledge management: A partnership between human and AI", *Business Horizons*, Vol. 66, No. 1, pp. 87–99.

Kelley, S. (2022) "Employee Perceptions of the Effective Adoption of AI Principles", *Journal of Business Ethics*, Vol. 178, pp. 871–893.

Kucharska, W. and Erickson, G.S. (2023) "Tacit knowledge acquisition & sharing, and its influence on innovations: A Polish/US cross-country study," *International Journal of Information Management*, 71, p. 102647.

Leoni, L., Gueli, G., Ardolino, M., Panizzon, M. and Gupta, S. (2024) "AI-Empowered KM Processes for Decision-Making: Empirical Evidence from Worldwide Organisations", *Journal of Knowledge Management*, Vol. 28, No. 11, pp. 320–347.

Li, L. (2022) "Reskilling and Upskilling the Future-ready Workforce for Industry 4.0 and Beyond", *Information Systems Frontiers*, Vol. 26, No. 5, pp. 1697–1712.

Lin, X. and Wu, D. (2025) "AI Technology Adoption, Knowledge Sharing, and Manufacturing Firms' Innovation Performance: The Moderating Effect of Absorptive Capacity", *IEEE Transactions on Engineering Management*, Vol. 72, pp. 2137–2149.

Mikalef, P., Lemmer, K., Schaefer, C., Ylinen, M., Fjørtoft, S.O., Torvatn, H.Y., … and Niehaves, B. (2023) "Examining How AI Capabilities Can Foster Organizational Performance in Public Organizations", *Government Information Quarterly*, Vol. 40, No. 2, 101797.

Mosqueira-Rey, E., Hernández-Pereira, E., Alonso-Ríos, D., Bobes-Bascarán, J. and Fernández-Leal, Á. (2022) "Human-in-the-loop machine learning: a state of the art", *AI Review*, Vol. 56, pp. 3005–3054.

Nair, M., Svedberg, P., Larsson, I. and Nygren, J. (2024) "A comprehensive overview of barriers and strategies for AI implementation in healthcare: Mixed-method design", *PLOS ONE*, Vol. 19.

Nawaz, A., Zhong, N., Ahsan, A. and Tian, R. (2025) "Organizational Learning in Projects: An Innovational Approach", *Journal of Posthumanism*, Vol. 5, No. 1, pp. 1354–1366.

Rostamzadeh, R., Alizadeh, F.K., Keivani, S. and Isavi, H. (2025) "The Role of AIin Improving Organizational Behavior: A Systematic Study", *Human Behavior and Emerging Technologies*, Vol. 2025, No. 1, 8094428.

Schwaeke, J., Gerlich, C., Nguyen, H. L., Kanbach, D. K., and Gast, J. (2025) "Artificial intelligence (AI) for good? Enabling organizational change towards sustainability", *Review of managerial science*, *19*(10), 3013-3038.

Storey, V. (2025) "Knowledge Management in a World of Generative AI: Impact and Implications", *ACM Transactions on Management Information Systems*, Vol. 16, pp. 1–14.

Tambe, P. (2025) “Reskilling the Workforce for AI: Domain Expertise and Algorithmic Literacy”, *Management Science*, Vol. 72, No. 1, pp. 515–537.

Tukey, J.W. (1949) “Comparing individual means in the analysis of variance”, *Biometrics*, 5(2), pp. 99–114.

Yan, J., Husted, K. and Fath, B. (2025) “Transforming Organizational Knowledge Creation Through Artificial Intelligence: A Systematic Review of the Emergent Literature”, *VINE Journal of Information and Knowledge Management Systems*, pp. 1–19.

Zhang, Q., Zuo, J., and Yang, S. (2025) “Research on the impact of generative artificial intelligence (GenAI) on enterprise innovation performance: a knowledge management perspective”, *Journal of Knowledge Management*, *29*(7), 2238-2257.

When Knowledge Hiding Enhances Self-Assessment in Multiple Sclerosis

Juan-Gabriel Cegarra-Navarro[1], Aurora Martínez-Martínez[1], Scott Erickson[2], Jorge Cegarra-Sánchez[3]
[1] Technical University of Cartagena, Spain
[2] Ithaca College, New York, United States
[3] University Hospital of Gran Canaria Dr. Negrín, Spain

Abstract

Small variations in how information is framed or selectively presented can significantly shape how individuals interpret and internalize knowledge. Prior research has largely associated knowledge hiding with negative outcomes, partly due to its operationalization as perceived lack of help or misaligned information. In contrast, this study adopts a knowledge regulation perspective, focusing on how selective non-disclosure may shape cognitive processes. Therefore, we adopt a complementary perspective by examining evasive knowledge hiding as a form of knowledge regulation, where information is filtered and sequenced to reduce cognitive overload and prevent the formation of misconceptions.

This study examines how evasive knowledge hiding influences self-assessment among patients with Multiple Sclerosis, a context in which misconceptions and narrow mental models can significantly affect disease management. We argue that learning myopia, understood as the persistence of distorted or incomplete mental models, plays a central role in shaping how patients evaluate their own capabilities and progress. Drawing on survey data from 75 Multiple Sclerosis patients and analysed using Partial Least Squares Structural Equation Modeling (PLS-SEM), the results reveal that evasive knowledge hiding does not directly affect self-assessment. Instead, its impact is fully mediated by learning myopia. Greater evasive knowledge hiding is associated with lower levels of learning myopia, which in turn leads to higher levels of self-assessment. These findings suggest that selective non-disclosure, when enacted as a form of knowledge regulation, may help patients avoid misleading or overly complex information that could distort their understanding.

By challenging the dominant "dark side" view of knowledge hiding, this study highlights its contingent and potentially adaptive role in specific contexts. The findings contribute to knowledge management research by reframing evasive knowledge hiding as a mechanism of knowledge regulation and to healthcare research by showing how managing information flows can support more accurate self-evaluations among patients. Overall, the study uncovers a paradoxical pathway through which knowledge hiding can enhance learning outcomes by reducing cognitive distortions.

Keywords: Evasive knowledge hiding; learning myopia; self-assessment; systems thinking; healthcare knowledge management.

Paper type: Academic

1 Introduction

Self-assessment is a central construct for understanding patient evolution because it marks the point at which knowledge and beliefs are transformed into action (Zheng et al., 2024). For individuals living with Multiple Sclerosis (MS), accurate self-assessment is not merely a psychological disposition but a practical capacity that guides everyday decisions, supports treatment adherence, and shapes interactions with healthcare providers (Tacchino et al., 2018). In contrast to broader constructs such as self-esteem, self-assessment focuses on situational judgments of one's abilities, confidence in managing challenges, and perceptions of progress in coping with the condition (Zell & Johansson, 2024).

From the perspective of individual learning, self-assessment can be hidden by learning myopia when misconceptions and narrow mental models distort how individuals interpret their own progress and capabilities (Dunning et al., 2004). Patients with MS who believe, for example, that the illness is inevitably fatal or that it only affects specific groups, are likely to undervalue their ability to manage the condition effectively (Cegarra-Sánchez et al., 2025). In this way, learning myopia clouds self-assessment, leading to either an exaggerated sense of limitation or an inaccurate sense of progress (Cegarra-Sánchez et al., 2025).

By reducing cognitive blind spots and encouraging broader, system-oriented thinking, healthcare providers can help patients form more balanced self-assessments, which are essential for informed decision-making and adaptive coping (Meissner et al., 2017; Pina e Cunha & Chia, 2007). The key question is how practitioners can achieve this in everyday encounters. From our perspective, evasive knowledge hiding offers one possible mechanism (Xia & Li, 2023). By filtering and sequencing information, providers prevent the reinforcement of

misconceptions and avoid overwhelming patients with details that may foster distorted interpretations (Hernaus et al., 2019). Seen through a systems thinking lens, such selective withholding is not a form of deception but a knowledge regulation strategy that maintains cognitive clarity. In this way, practitioners can support patients in moving beyond learning myopia, enabling more realistic and empowering self-assessments.

Despite the growing research on knowledge hiding, most studies describe it as a detrimental behavior with negative consequences for collaboration and learning (Černe et al., 2014; Connelly et al., 2012; Zhang & Min, 2021). Less attention has been paid to its potential adaptive functions (Cegarra-Navarro et al., 2021), particularly in healthcare settings where patients are vulnerable to misconceptions (Zhao & Xia, 2018). Building on this gap, our study explores the paradoxical role of evasive knowledge hiding in supporting patient learning. Specifically, we address the following research questions:

- *RQ1: How does evasive knowledge hiding influence cognitive myopia among patients with Multiple Sclerosis?*
- *RQ2: What is the relationship between cognitive myopia and self-assessment in this context?*
- *RQ3: Does cognitive myopia mediate the relationship between evasive knowledge hiding and self-assessment?*

The remainder of the paper is organized as follows. Section 2 develops the conceptual framework and hypotheses, drawing on the literatures on knowledge hiding, learning myopia, and systems thinking. Section 3 describes the methodology, including the sample, measures, and analytical procedures. Section 4 presents the results of the PLS-SEM analysis. Section 5 discusses the findings, highlighting their theoretical and practical implications. Finally, Section 6 concludes with limitations and directions for future research. In doing so, this study reframes knowledge hiding as a knowledge regulation strategy and demonstrates its paradoxical role in enhancing patient self-assessment.

2 Conceptual Framework

From a knowledge management perspective, self-assessment is not only a psychological trait but a knowledge outcome that reflects whether information has been effectively absorbed, contextualized, and applied (Xu et al., 2010). Nonaka and Takeuchi (1995) highlight that knowledge becomes valuable when it is internalized and externalized in practice, in this vein, self-assessment captures both dimensions by showing how patients internalize information and communicate their evaluation of capabilities. Similarly, Argyris and Schön (1996) emphasize reflection as a cornerstone of learning, where individuals reassess assumptions to adapt their actions. Davenport et al. (1998) similarly argue that knowledge emerges when information is enriched by reflection and experience, which positions self-assessment as a practical indicator of knowledge quality and effectiveness.

What these ideas underscore is that accurate self-assessment reflects the successful transformation of knowledge into actionable understanding, while distorted self-assessment reveals breakdowns in knowledge flows. In this sense, self-assessment provides a valuable lens for knowledge management research, as it not only signals whether knowledge is available but also whether it has been meaningfully applied in practice (Xu et al., 2010). This perspective becomes especially relevant in the case of MS, where learning myopia often shapes patient experience. Learning myopia refers to persistent misconceptions and narrow mental models, such as the belief that “MS affects only women” or that “MS is inherently fatal” (Cegarra-Sánchez et al., 2025).

Systems thinking views healthcare interactions as knowledge ecosystems. Within such systems, too much, too little, or poorly timed information can either destabilize or stabilise learning (Adam & De Savigny, 2012; Cresswell et al., 2021). From this perspective, learning myopia and systems thinking can be viewed as two sides of the same coin, whereas systems thinking emphasizes interconnections, feedback, and long-term implications (Thelen et al., 2023), learning myopia represents fragmented and short-sighted processing of information (Cegarra-Sánchez et al., 2025).

In healthcare contexts, the absence of systems thinking may trap patients in cognitive myopia, reinforcing misconceptions and limiting their self-assessment (Gao & Yin, 2025). Conversely, when providers adopt a systems thinking perspective, they are able to filter, organize, and sequence information in ways that reduce misperceptions and promote more adaptive learning (Adam & De Savigny, 2012). It should be noted here that organizational research has shown that hiding behaviors influence what becomes salient and shared in groups, with climates, leadership signals, and communication patterns directing attention and shaping belief formation (Ul Haq et al., 2024; C. Xu & Yao, 2025; Yang et al., 2024).

From a systems thinking perspective, evasive knowledge hiding (EKH) should not be understood simply as withholding knowledge. Although it shares similarities with information filtering and communication tailoring, EKH is distinct in that it involves intentional non-disclosure under conditions of asymmetric knowledge, where

clinicians control access to information that patients may expect or request. It also reflects a temporary and situational process, in which information is not necessarily denied but selectively delayed or redirected. These features position EKH within the broader knowledge hiding literature rather than as a purely neutral or patient-centered communication practice.

At the same time, EKH can be seen as a form of information filtering that regulates knowledge flows, helping patients process information without becoming overwhelmed or misled by distorted messages. While much of the literature associates knowledge hiding with negative antecedents such as the dark triad (Joshi et al., 2025), abusive contexts and organizational politics (Wang et al., 2024), or identity threats (Kumar et al., 2025; Lu & Zhou, 2025), and links it to harmful outcomes for collaboration and performance (Zhang & Min, 2024), more recent research highlights its contingent nature. In particular, supportive climates, leadership signals, and relational contexts can buffer or redirect its effects, revealing that knowledge hiding may not always be detrimental (Chacon-Henao & Arias-Perez, 2025; Fu et al., 2024; Islam et al., 2024; Semerci et al., 2025; C. Xu & Yao, 2025; Yang et al., 2024).

It follows from the above that the strength of the effect of evasive knowledge hiding on self-assessment may vary across contexts (Islam et al., 2024; C. Xu & Yao, 2025). Supportive communication climates, relational cues such as identity leadership or gratitude, and patient characteristics such as age or health literacy can amplify the benefits of selective withholding by sustaining trust and reducing susceptibility to misconceptions (Chacon-Henao & Arias-Perez, 2025; C. Xu & Yao, 2025). In contrast, when hiding arises as a self-protective response to stress, or in environments marked by moral disengagement and politics, evasive knowledge hiding may lose its adaptive function and reinforce rather than reduce misconceptions (Kumar et al., 2025; Wang et al., 2024). Importantly, this interpretation assumes that EKH is enacted with patient-centered intent and within a context of trust. When these conditions are absent, selective withholding may be perceived as neglect or paternalism, potentially undermining its beneficial effects.

Translating these ideas into clinical encounters, practitioners usually aim to serve the best interests of patients while managing cognitive overload, uncertainty, and time pressures. In such settings, selectively withholding details that are confusing or of little relevance can shield patients from misinformation, misattribution, or unnecessary rumination, particularly when their understanding of the illness is fragile. In this sense, evasive knowledge hiding can act as protective filtering rather than opportunistic concealment. By limiting exposure to misleading lay beliefs and preventing simplistic but inaccurate explanations from taking hold, clinicians reduce the likelihood that patients interpret their condition through misconceptions. This process lowers cognitive myopia. Research in knowledge management supports this logic, showing that environments which minimize noise and conflict improve clarity and strengthen self-regulatory capacity, whereas toxic climates erode it (Lu & Zhou, 2025; Mohamed et al., 2024; Wang et al., 2025). For patients with Multiple Sclerosis, fewer misconceptions lead to clearer mental models and more accurate self-evaluations of capability and progress. Therefore, we propose:

H1. *Evasive knowledge hiding is negatively related to learning myopia*.

Learning myopia narrows the cognitive space through which individuals evaluate themselves. When misconceptions persist, patients judge their progress and abilities through distorted frames, which can lead either to excessive pessimism or to an inaccurate sense of improvement. In this way, learning myopia undermines the capacity for realistic self-assessment. Evidence from organizational research shows that hidden or distorted knowledge often reduces performance because it erodes trust, generates conflict, and restricts opportunities for learning (Islam et al., 2024; Zhang & Min, 2024). In healthcare, this mechanism is primarily cognitive. Patients who remain trapped in narrow or mistaken beliefs about their illness lack the clarity needed to assess their situation effectively. By contrast, when learning myopia is reduced, patients are more likely to develop coherent mental models and provide accurate evaluations of their ability to cope. Therefore, we propose the second hypothesis:

H2. *Learning myopia is negatively related to self-assessment*.

This study conceptualises evasive knowledge hiding (EKH) as the selective non-disclosure or deflection of information that a provider judges as unnecessary, untimely, or potentially confusing in a specific interaction. Unlike rationalised hiding or playing dumb, EKH is not primarily hostile or competitive. It represents a situational choice to limit detail, postpone explanations, or steer away from topics considered irrelevant at that moment. This nuance differentiates EKH from the broader family of deceptive or self-protective hiding described in organisational settings, where intentions often relate to impression management, politics, or resource protection (Bashir et al., 2024; Wang et al., 2024; Wiggins, 2024; Xu et al., 2024).

As discussed previously, the effect of EKH on self-assessment is likely to depend on the context. Importantly, when knowledge is selectively withheld in a constructive way, it should not exert a direct influence on self-assessment. Instead, the main mechanism operates through cognition. By filtering and sequencing information,

clinicians help patients avoid misconceptions that would otherwise distort their judgments (Jones, 2024). This process reduces learning myopia, and as a result, patients develop clearer mental models and are better able to evaluate their own capabilities and progress. Thus, the mediation pathway explains why the direct link between EKH and self-assessment may be weak, while the indirect effect through learning myopia is expected to be positive.

H3. *Evasive knowledge hiding has a positive indirect effect on self-assessment through learning myopia*.

3 Method

3.1 Data collection

The study was carried out at the individual level, focusing on people living with Multiple Sclerosis (MS). Participants were between 22 and 67 years old, with 32 men and 51 women included in the sample. The wide age span was intentional, since age can shape how patients perceive and respond to their condition. Data were collected in October and November 2024 through a structured survey distributed by Multiple Sclerosis Associations, resulting in 75 valid responses. The questionnaire also captured demographic characteristics such as participants' region and province of residence.

To ensure sample adequacy, a power analysis was performed with G*Power (Cunningham & McCrum-Gardner, 2007). Following Cohen's (1998) recommendations, where f^2 values of 0.02, 0.15, and 0.35 represent small, medium, and large effects, the minimum requirement for a medium effect size with a one-tailed test was 74 respondents. The final sample of 75 thus satisfied this criterion.

Potential non-response bias was examined by comparing Self-Assessment scores from early and late respondents using an independent samples t-test. The results showed no significant differences ($p = 0.807$). Common method bias (CMB) was assessed following Podsakoff et al. (2012). Harman's single-factor test indicated that the first factor accounted for 40.38% of the variance, below the 50% threshold (Podsakoff et al., 2012). In addition, a confirmatory factor analysis (CFA) demonstrated that a one-factor model did not fit the data well (Satorra–Bentler $\chi^2(36) = 122.265$, $\chi^2/df = 3.39$) compared with the measurement model ($\chi^2(24) = 31.420$, $\chi^2/df = 1.30$). Taken together, these analyses confirmed that CMB was not a major concern (Armstrong & Overton, 1977).

3.2 Measures

In this study, Evasive Knowledge Hiding (EKH) was modeled as a reflective construct using three items adapted from (Connelly et al., 2019). These items capture patients' perceptions of whether medical staff avoided sharing information or support when it was requested. EKH emphasizes situations in which help was delayed, information was withheld, or guidance proved unhelpful. Importantly, this form of knowledge hiding may involve concealing information that is perceived as irrelevant by the provider, rather than intentionally withholding essential knowledge.

Cognitive Myopia (CM) was measured using three items adapted from (Cegarra-Sánchez et al., 2025). This construct represents the persistence of misconceptions or oversimplified beliefs about Multiple Sclerosis. The selected items capture distorted perceptions of the illness, for example, viewing it as exclusive to women, inherently fatal, or the result of unhealthy lifestyle choices. Together, they reflect a narrowed and biased understanding of the disease, consistent with the idea of cognitive blind spots.

Finally, Self-Assessment (SA) was measured with three reflective indicators based on (Fung et al., 2020). These items focus on individuals' ability to evaluate their own capabilities and progress. Respondents were asked to assess whether they feel as capable as others in managing their condition, whether they are confident in their ability to cope with challenges, and whether they maintain a positive outlook on their daily progress. These items align with prior research on self-concept and self-efficacy, while emphasizing the role of reflective self-evaluation in the context of chronic illness.

3.3 Assessment of the measures

Given the reflective nature of the constructs, all variables were modeled as Type A composites, where the direction of causality flows from the latent factor to its indicators Cepeda-Carrion et al., (2019). Following the guidelines of Henseler et al. (2016), we assessed item reliability, composite reliability, and convergent validity. All indicators loaded strongly on their constructs, with values ranging from 0.600 to 0.906. One item (CM1) fell slightly below the 0.70 threshold, loading at 0.600. We decided to retain it because it reflects an important aspect of Cognitive Myopia, is consistent with prior definitions of the construct, and its removal did not improve

reliability or validity. This decision aligns with established guidelines, which allow the inclusion of items with loadings between 0.50 and 0.70 when supported by theory and when the overall measurement quality remains strong (Hair et al., 2019; Henseler et al., 2016). Reliability and validity were confirmed across all constructs. Composite reliability values were above 0.80, and AVE scores exceeded 0.50 (EKH: AVE = 0.663, ρA = 0.855; CM: AVE = 0.616, ρA = 0.824; SA: AVE = 0.695, ρA = 0.872). Variance inflation factors ranged from 1.226 to 2.310, far below the cutoff of 5, ruling out multicollinearity concerns. Overall, the results confirm that the three constructs (i.e. Evasive Knowledge Hiding, Cognitive Myopia, and Self-Assessment), were measured with good reliability and validity.

Discriminant validity was assessed using the Fornell–Larcker criterion and the Heterotrait–Monotrait ratio (HTMT). In doing so, the square roots of the AVE values, shown on the diagonal, were higher than the correlations among the constructs, which satisfies the Fornell–Larcker criterion. In addition, all HTMT values were below the conservative threshold of 1.0 (Henseler et al., 2014), with the highest observed value being 0.818. These findings confirm that the constructs Evasive Knowledge Hiding (EKH), Cognitive Myopia (CM), and Self-Assessment (SA) are empirically distinct, supporting discriminant validity.

4 Results

Table 1 presents the structural model results. The results indicate a case of full mediation, as the direct effect of Evasive Knowledge Hiding (EKH) on Self-Assessment (SA) is not significant (β = -0.304, p = 0.135), while the indirect effect through Cognitive Myopia (CM) is significant (β = 0.210, p = 0.019). This suggests that EKH influences SA exclusively through its effect on CM, confirming the presence of total mediation (Hair et al., 2022; Zhao et al., 2010). These results support H1 and H2 and confirm the indirect effect proposed in H3.

Table 1: Effects on endogenous variables

Direct effects	Point estimate	Confidence Intervals 2.5%	97.5%	p-values	R^2	f^2
H_1: EKH→CM	-0.355	-0.646	-0.247	0.001	0.126	0.144
H_2: CM→SA	-0.707	-0.766	-0.234	0.000	0.429	0.585
H_3: EKH→SA	-0.304	-0.122	0.371	0.135	0.429	0.039
Indirect effects						
EKH→CM→SA	0.251	0.085	0.390	0.009	0.429	0.039

Global goodness of fit, Confirmatory composite analysis and bootstrap-based 95% and 99% quantiles

Goodness of Fit Measure	Estimated Model	Hi95	Hi99	Saturated Model	Hi95	Hi99
SRMR	0.083	0.112	0.132	0.083	0.112	0.132
d_{ULS}	0.533	0.563	0.782	0.163	0.563	0.782
d_G	0.224	0.328	0.436	0.056	0.328	0.436

Notes: Evasive Knowledge Hiding→(EKH); Cognitive Myopia →(CM); *Self-Assessment*→(SA).

5 Discussion

This study set out to examine the paradoxical role of evasive knowledge hiding in shaping patient learning and adaptation in the context of Multiple Sclerosis. The results confirm that evasive knowledge hiding reduces learning myopia (H1), learning myopia negatively affects self-assessment (H2), and that the effect of evasive knowledge hiding on self-assessment is transmitted indirectly through learning myopia (H3). The direct path from evasive knowledge hiding to self-assessment was non-significant, confirming that its influence is primarily cognitive and operates through the reduction of misconceptions.

Our findings contribute to a growing body of literature that challenges the traditional view of knowledge hiding as exclusively harmful. Most studies describe knowledge hiding as a dysfunctional behavior linked to distrust, conflict, and reduced performance (Černe et al., 2014; Connelly & Zweig, 2015; Zhang & Min, 2021). Recent work, however, has begun to explore its paradoxical aspects, noting that under certain conditions, selective withholding can protect knowledge quality and preserve group functioning (Chacon-Henao & Arias-Perez, 2025; Restrepo-Gil et al., 2024). The current study extends this line of inquiry by showing that in healthcare encounters, evasive knowledge hiding can operate as a form of protective filtering. By reducing patients' exposure to misconceptions, clinicians create conditions that limit cognitive distortions and support more accurate self-assessments.

The findings also align with prior research on learning myopia, which describes the persistence of narrow or mistaken cognitive frames that restrict adaptive learning (Cegarra-Sánchez et al., 2025). Consistent with

organizational studies showing that hidden or distorted knowledge reduces performance by undermining clarity (Islam et al., 2024; Zhang & Min, 2024), our results show that learning myopia has a detrimental impact on patients' ability to evaluate their own progress and capabilities. Importantly, our study bridges this construct with self-assessment, a psychological but also knowledge-based outcome that reflects the transformation of information into action (Nonaka & Takeuchi, 1995; Y. Xu et al., 2010).

Regarding H3, the mediation result is consistent with a systems thinking perspective. Systems thinking emphasizes the regulation of knowledge flows to maintain coherence and avoid overload (Adam & De Savigny, 2012). From this lens, evasive knowledge hiding is not a form of deception but a filtering mechanism that stabilizes patient cognition by reducing noise and limiting misperceptions.

6 Conclusion

This study makes three main contributions. First, it reframes evasive knowledge hiding as a knowledge regulation mechanism, extending the dominant "dark side" view. Second, it introduces learning myopia as a cognitive mechanism linking knowledge processes to self-assessment. Third, it integrates systems thinking into knowledge hiding research by showing how regulating information flows shapes cognitive outcomes.

Methodologically, the study applies PLS-SEM with endogeneity checks to a novel dataset of MS patients, extending the application of KM research methods into clinical settings. Conceptually, it demonstrates the value of focusing on indirect effects and mediation pathways, providing a more nuanced understanding of how knowledge processes translate into patient outcomes.

Practically, the findings suggest that healthcare providers can improve patient adaptation not by sharing every piece of information, but by carefully regulating knowledge flows. Selective withholding of low-value or confusing details can prevent misconceptions and enable patients to build clearer self-assessments, which are critical for adherence and long-term coping. This perspective has implications for medical training, communication strategies, and the design of patient education programs.

The study has limitations that open avenues for further research. First, the sample size, while adequate for PLS-SEM, was relatively small and limited to one national context. Future studies should test the model across larger and more diverse populations to enhance generalizability. Second, the cross-sectional design prevents causal inference; longitudinal or experimental studies could validate the mediation pathway over time. Third, while we focused on evasive knowledge hiding, future work could explore how other forms of knowledge hiding (rationalized, playing dumb) operate in clinical settings, and whether they have different or even harmful effects on patient learning. Finally, future research should examine moderators such as communication climate, leadership cues, or patient characteristics (e.g., age, health literacy) that may strengthen or weaken the positive role of evasive knowledge hiding.

7 Ethics Declaration

This study was conducted using anonymized survey data collected through Multiple Sclerosis Associations publicly available. No sensitive personal data were recorded, and participation was voluntary. According to applicable regulations, formal ethical approval was not required.

8 AI Declaration

AI tools were used to support language refinement and clarity of expression. No AI tools were used in the data analysis, interpretation of results, or development of the theoretical framework. The authors take full responsibility for the content of the paper.

References

Adam, T., & De Savigny, D. (2012). Systems thinking for strengthening health systems in LMICs: need for a paradigm shift. Health Policy and Planning, 27 Suppl 4. https://doi.org/10.1093/HEAPOL/CZS084

Bashir, H., Fanchen, M., & Bari, M. W. (2024). Deceptive Knowledge Hiding in Organizations: Psychological Distress as an Underlying Mechanism. SAGE OPEN, 14(2). https://doi.org/10.1177/21582440241251996

Cegarra-Sánchez, J., Cegarra-Navarro, J.-G., & Martinez-Martinez, A. (2025). Addressing Patient Learning Myopia in Multiple Sclerosis Through A Thermodynamic Approach. Journal of Nursing Care Quality. https://journals.lww.com/jncqjournal/fulltext/9900/addressing_patient_learning_myopia_in_multiple.199.aspx

Chacon-Henao, J., & Arias-Perez, J. (2025). Counteracting knowledge hiding among academics to foster creativity: is organisational support the silver bullet or fools' gold? VINE JOURNAL OF INFORMATION AND KNOWLEDGE MANAGEMENT SYSTEMS. https://doi.org/10.1108/VJIKMS-04-2023-0091

Cresswell, K., Sheikh, A., Franklin, B., Krasuska, M., Nguyen, H. T., Hinder, S., Lane, W., Mozaffar, H., Mason, K., Eason, S., Potts, H., & Williams, R. (2021). Interorganizational Knowledge Sharing to Establish Digital Health Learning Ecosystems: Qualitative Evaluation of a National Digital Health Transformation Program in England. Journal of Medical Internet Research, 23. https://doi.org/10.2196/23372

Dunning, D., Heath, C., & Suls, J. (2004). Flawed Self-Assessment. Psychological Science in the Public Interest, 5, 106–169. https://doi.org/10.1111/j.1529-1006.2004.00018.x

Gao, L., & Yin, H. (2025). Managerial myopia and greenwashing: A machine learning and text analysis approach. INTERNATIONAL REVIEW OF FINANCIAL ANALYSIS, 105. https://doi.org/10.1016/j.irfa.2025.104452

Hernaus, T., Cerne, M., Connelly, C., Poloski Vokic, N., & Škerlavaj, M. (2019). Evasive knowledge hiding in academia: when competitive individuals are asked to collaborate. Journal of Knowledge Management, 23(4), 597–618. https://doi.org/10.1108/JKM-11-2017-0531

Meissner, P., Brands, C., & Wulf, T. (2017). Quantifiying blind spots and weak signals in executive judgment: A structured integration of expert judgment into the scenario development process. International Journal of Forecasting, 33(1), 244–253. https://doi.org/10.1016/j.ijforecast.2015.08.002

Restrepo-Gil, D., Suarez-Castellanos, O., & Arias-Perez, J. (2024). The paradox of knowledge hiding among academics: Implications for intellectual capital-driven knowledge transfer. INDUSTRY AND HIGHER EDUCATION. https://doi.org/10.1177/09504222241299670

Tacchino, A., Brichetto, G., Zaratin, P., Battaglia, M., & Ponzio, M. (2018). Self-assessment reliability in multiple sclerosis: the role of socio-demographic, clinical, and quality of life aspects. Neurological Sciences, 40, 617–620. https://doi.org/10.1007/s10072-018-3589-6

Thelen, J., Fruchtman, S., Bilal, M., Gabaake, K., Iqbal, S., Keakabetse, T., Kwamie, A., Mokalake, E., Mupara, L., Seitio-Kgokgwe, O., Zafar, S., & Muñoz, C. (2023). Development of the Systems Thinking for Health Actions framework: a literature review and a case study. BMJ Global Health, 8. https://doi.org/10.1136/bmjgh-2022-010191

Ul Haq, I., De Clercq, D., & Azeem, M. U. (2024). How and when do perceptions of supervisor evasive knowledge hiding escalate into diminished job performance? CANADIAN JOURNAL OF ADMINISTRATIVE SCIENCES-REVUE CANADIENNE DES SCIENCES DE L ADMINISTRATION, 41(4), 558–576. https://doi.org/10.1002/cjas.1745

Wang, Y., Bari, M. W., Shaheen, S., & Liu, X. (2024). Dark triad and knowledge hiding in firms: Mediating role of perceived organizational politics. ACTA PSYCHOLOGICA, 248. https://doi.org/10.1016/j.actpsy.2024.104339

Xu, C., & Yao, Z. (2025). Exploring the effect of line manager gratitude expression on employees' knowledge hiding and unethical pro-supervisor behaviors. JOURNAL OF KNOWLEDGE MANAGEMENT, 29(6), 1951–1973. https://doi.org/10.1108/JKM-02-2024-0127

Xu, Y., Zhan, H. F., Yu, J. H., & Sun, L. (2010). Design of a Knowledge-Based Self-Assessment System for Enterprise Knowledge Management Evaluation. Applied Mechanics and Materials, 37–38, 322–326. https://doi.org/10.4028/www.scientific.net/AMM.37-38.322

Zell, E., & Johansson, J. (2024). The Association of Self-Esteem With Health and Well-Being: A Quantitative Synthesis of 40 Meta-Analyses. Social Psychological and Personality Science. https://doi.org/10.1177/19485506241229308

Zhang, Z., & Min, M. (2021). Organizational rewards and knowledge hiding: task attributes as contingencies. Management Decision, 71902107. https://doi.org/10.1108/MD-02-2020-0150

Zheng, B., He, Q., & Lei, J. (2024). Informing factors and outcomes of self-assessment practices in medical education: a systematic review. Annals of Medicine, 56 1, 2421441. https://doi.org/10.1080/07853890.2024.2421441

Beyond Single Tasks: How Students Interact with AI in Design Processes

Matteo Meacci, Antonella Martini, Gabriele Goretti, Caterina Dastoli
University of Pisa, Pisa, Italy

Abstract

The advent of Generative Artificial Intelligence (GenAI) has led to a rapid expansion of AI use in educational contexts. This new technology is profoundly transforming students' learning experiences. However, a limitation of current research is its predominant focus on how AI influences student performance, concentrating on quantitative indicators such as test scores. In contrast, research on how AI tools are perceived and used by students remains limited, and it is still unclear how collaboration with AI supports students' knowledge-building processes. This study addresses this gap by exploring how students interact with AI and how they feel during these interactions in design processes. A mixed-methods study was conducted to investigate this question, involving 16 students enrolled in a bachelor's degree program in Industrial Design Engineering. The students were divided into groups and participated in a two-day workshop, during which they worked on product development following the Double Diamond process and supported by KOBI, an AI tool specialized in the field of arts and design. During the workshop, conversations between the students and the AI tool were recorded, along with data collected through questionnaires administered before and during the workshop activities. The data were analyzed through descriptive statistics and qualitative content analysis. The results show that students interact with AI in different ways depending on the cognitive nature of the task. Two main modes of interaction were identified: exploratory and executive. These modes differ in both the form of interaction and the type of request addressed to AI. The findings also show that these different modes of AI use influence students' perceptions of the tool's usefulness and lead to different emotional experiences. Overall, the study contributes to the literature on generative AI in education by deepening understanding of how students interact with AI systems. Furthermore, it proposes an interpretive model, inspired by the concept of absorptive capacity, that links the identified modes of AI use to students' knowledge-building processes.

Keywords: Generative Artificial Intelligence in Education, Human-AI interaction, Students' Perceptions of AI, Mixed-methods Research, Design Process.

Paper type: Academic Research Paper

1 Introduction

The application of artificial intelligence in education (AIEd) has been the subject of research for about thirty years (Zawacki-Richter et al., 2019). However, it is with the advent of Generative Artificial Intelligence (GenAI), which officially began with the release of ChatGPT in November 2022, that we have seen a rapid spread of AI use in educational contexts (Lodge, 2023; Batista et al., 2024; Yusuf et al., 2024). This new technology has emerged as an innovative force capable of transforming education, promising to personalize learning experiences by adapting courses to individual progress, increase student engagement, automate administrative tasks, reduce workloads, and enhance teacher skills (Kasneci et al., 2023; Al-Zahrani & Alasmari, 2024; Batista et al., 2024; Sahar & Munawaroh, 2025).

Despite these opportunities, the introduction of AI in educational contexts also raises numerous ethical and social challenges and concerns, including issues related to data privacy and algorithmic bias (Zawacki-Richter et al., 2019; Al-Zahrani & Alasmari, 2024; Batista et al., 2024; Sahar & Munawaroh, 2025). Academic integrity is also a challenge that is fueled by fears that students' cognitive and logical skills could be reduced by excessive use of generative AI tools (Yusuf et al., 2024).

One limitation of research on AIEd is that, in most cases, it has focused almost exclusively on the effectiveness of AI tools in improving individual students' academic performance on specific individual tasks. In particular, attention has been focused primarily on quantitative indicators such as improved test scores (Holmes & Tuomi, 2022; Muzumdar & Cheemalapati, 2025; Pulungan & Nora, 2025).

However, limiting the analysis to performance outcomes risks overlooking a more profound transformation. Generative AI does not simply act as a tool that improves task efficiency but participates in the dynamics of knowledge construction (Chen et al., 2023; Cress & Kimmerle, 2023). AI systems can support how knowledge is created, shared, and applied within learning environments (Tsoi & Strønen, 2024; Hagos et al., 2025). Rather than simply supporting individual outcomes, AI can reshape how students construct meaning, explore alternatives, and externalize ideas (Ahmed, 2025; Muzumdar & Cheemalapati, 2025; Pulungan & Nora, 2025).

This requires moving beyond outcome-based assessments toward a process-oriented understanding of student-AI interaction. In this regard, research on how AI tools are perceived and used by students remains limited, particularly on what students feel, how they reflect, and how they attribute meaning to what the AI tool does (Muzumdar & Cheemalapati, 2025; Pulungan & Nora, 2025). These emotional and reflective components of the experience play a crucial role in learning. We still know little about how collaboration with AI tools helps students build knowledge. The significance of this gap lies in the need to know how students interact with AI tools to ensure their effective integration into educational contexts (Strzelecki, 2024; Muzumdar & Cheemalapati, 2025).

This study addresses the above gap by exploring the following research question: How do students interact with AI tools, what do they ask, and how do they feel when interacting with AI in design processes?

Design processes refer to a structured sequence of interdependent activities, distinct from specific individual tasks. Particularly, design activities have long been recognized as inherently iterative, reflective, and meaning-making processes in which tools actively shape thinking and collaboration (Schön, 1983). This makes design a particularly relevant domain for examining how students perceive, engage with, and construct knowledge alongside generative AI, beyond measurable performance outcomes.

The study contributes to the literature on generative AI in education by deepening understanding of how students interact with AI and how they experience that interaction. Specifically, it shows that AI use is not uniform but changes depending on the cognitive demands of the task, triggering different modes of use that influence the student's experience. Furthermore, this study links the identified modes of AI use to the knowledge-building process and proposes an interpretive model based on the concept of absorptive capacity introduced by Cohen and Levinthal (1990).

The remainder of the paper is structured as follows. Section 2 describes the method adopted to explore the research question. Section 3 presents the results of the study. Section 4 discusses and interprets the findings, illustrates the proposed model, and outlines the main limitations of the study. Finally, Section 5 presents the conclusions.

2 Method

2.1 Study design and participants

To investigate the research question, an exploratory mixed-methods study was conducted involving 16 second-year students from the bachelor's degree course in Industrial Design Engineering. The students were divided into six groups and participated in a two-day workshop. During the workshop, they worked on product development starting from specific areas of reference (e.g., wood, marble, glass, etc.), following the Double Diamond process introduced by the Design Council in 2005 and supported by KOBI, an artificial intelligence tool specialized in the field of arts and design (Ripa di Meana et al., 2024).

2.2 Workshop setting

2.2.1 Double Diamond process

The workshop was structured according to the Double Diamond model (Design Council, 2005), a consolidated framework that splits the design process into divergent and convergent phases. The model consists of four phases:

- Discover, a divergent phase to identify the problem;
- Define, a convergent phase to frame the design challenge;
- Develop, a divergent phase to generate ideas;
- Deliver, a convergent phase to select and formalize the solution.

On the first day, the students completed the first two phases, while on the second day they completed the next two.

The Double Diamond was chosen for two main reasons. First, the structured alternation between moments of divergence and convergence allowed us to observe the students' interaction with the artificial intelligence tool in relation to cognitively different tasks. Second, participants were already familiar with this model, as it had previously been introduced in a university course and applied in a project carried out as part of the same course.

2.2.2 AI tool: KOBI

During the workshop, students were supported by KOBI, a digital knowledge ecosystem for research and education in the field of arts and design, developed by Ripa di Meana et al. (2024). Each group had access to one shared KOBI interface, which was collectively used during the design activities. KOBI uses Large Language Models (LLMs) to offer a nuanced understanding of the concepts required, drawing on a database of art and design literature.

The choice of KOBI over more mainstream tools (e.g., ChatGPT) was motivated by several characteristics. First, its specialization in the fields of arts and design. Secondly, in addition to a chat-based conversational interface, KOBI offers a three-dimensional representation of semantic space, allowing users to visualize the relationships between concepts related to the topic under discussion. This feature facilitates navigation between conceptual connections. Lastly, the system provides searchable references to support the information generated, allowing students to dive deeper into specific topics that emerge during interaction.

2.3 Data collection

During the workshop, conversations between students and the AI tool were recorded, for a total of 188 interactions, where each interaction corresponds to a prompt–response pair. Each group had a designated member responsible for systematically documenting the prompts entered and the corresponding responses generated by KOBI, organizing them according to the different phases of the Double Diamond. In addition, three questionnaires were administered to the participants. The first questionnaire was distributed before the beginning of the workshop to investigate students' familiarity with artificial intelligence tools and their perceptions of this technology. Two additional questionnaires were administered at the end of each workshop day. These questionnaires explored how students used AI during the activities, how they felt while interacting with it, and how useful they perceived the AI tool to be across the different phases of the Double Diamond. The first two aspects were investigated through open-ended questions, while perceived usefulness was measured on a 5-point rating scale.

2.4 Data analysis

2.4.1 Student-AI interaction analysis

The analysis of student–AI interactions was divided into three parts: a preliminary descriptive analysis, an analysis to identify students' intentions, and an analysis to highlight the themes of the requests. The term "themes" refers to what students asked the AI, while the term "intentions" refers to how those requests were formulated.

The analysis of students' intentions was conducted using a rule-based deductive content analysis, divided into two levels of coding applied to students' textual prompts.

The first level of coding concerns the form of interaction with the tool. A prompt is classified as a question if it contains a question mark or interrogative markers such as the 5Ws or How. If the prompt does not fall into this category but contains conversational request formulas, such as can you or could you, it is classified as a request, representing explicit requests that are not formulated in an imperative form. All remaining prompts are classified as commands, requests mainly formulated in an imperative form, in which the student asks the system to perform an action directly (e.g., generate, create, write).

The second level of coding concerns the type of request made to the tool. Specifically, requests were classified into two main categories: informational requests and generative requests. The classification was based on the identification of specific verbs and keywords. A prompt is classified as a generative request if it contains verbs that imply the production of structured content (e.g., write, generate, create) or project output objects (e.g., brief, scenario, personas). Conversely, a prompt is classified as an informational request if it contains words or expressions related to the search for information, such as information, examples, case studies. In cases where a prompt does not have sufficient indicators to be assigned to a specific category, it is classified as an unspecified request.

Subsequently, an analysis of the topics of the requests was conducted to identify the main thematic areas emerging in the interactions between students and AI. This analysis was carried out using inductive content analysis, to allow categories to emerge directly from the data, without imposing any prior thematic schemes.

The prompts were transformed into embeddings using a pre-trained language model for encoding textual meaning. The vectors obtained were then normalized to eliminate the effect of lexical redundancy and

expressive intensity. This operation made it possible to evaluate the similarity between prompts in terms of semantic affinity, making the Euclidean distance equivalent to the cosine similarity between vectors.

The categories were identified using an unsupervised density-based clustering algorithm (HDBSCAN), which identifies groups of semantically similar prompts only when there is sufficient density, avoiding the forced assignment of all prompts to a cluster. Because the algorithm does not assign semantic labels to the identified clusters, the interpretation and naming of the resulting categories was done manually by analyzing representative examples for each cluster. Prompts that were not automatically assigned to clusters or that were characterized by semantic ambiguity were subsequently examined and assigned to existing categories or, where necessary, used to define new categories.

In an initial implementation, the algorithm tended to group prompts mainly based on the starting contexts of the different groups (e.g., glass, wood, marble), rather than on the type of request. To remove this effect, semantic masking was introduced to neutralize the domain terms related to materials that influenced clustering. This made it possible to bring out themes related to the types of requests made to the AI, regardless of the specific project context.

The analyses were performed using Python scripts and then manually checked to verify the consistency of the results and refine the classification where necessary.

2.4.2 Questionnaire analysis

Data collected through questionnaires were analyzed using basic descriptive statistics and adopting the same approach used for analyzing interaction themes. Categories were identified from the open-ended responses regarding how students felt when using and interacting with AI during the different phases of the Double Diamond. Starting from these categories, a manual sentiment assessment was conducted by assigning +1 to positive connotations, -1 to negative ones, and 0 to neutral or unclassifiable categories. The values thus assigned were then aggregated for each phase, obtaining an overall sentiment score.

The results of the questionnaires were triangulated with those derived from the analysis of student-AI interactions to compare the usage patterns observed with the perception of usefulness and the experience reported by students during the design process.

3 Results

3.1 Two modes of student–AI interaction across the design process

The results show that the use of AI is not uniform across the design process, but changes in relation to the type of cognitive task that characterizes the phases. Different modes of interaction emerge, influencing the perception of usefulness and generating different emotional experiences. Two prevalent modes of interaction with AI have been identified: an exploratory mode and an executive mode. The transition from one mode to another is reflected both in the form of interaction and in the type of requests.

The results of the analysis of the intentions of student–AI conversations are shown in Figure 1 and Figure 2, which show the percentage distribution by phase of the types of requests and forms of interaction, respectively.

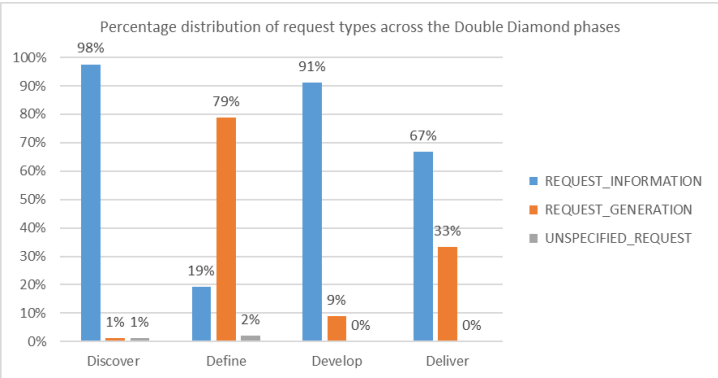


Figure 1: Percentage distribution of request types across the Double Diamond phases

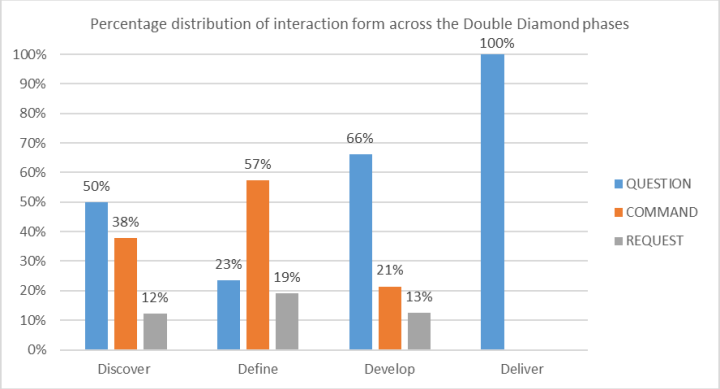


Figure 2: Percentage distribution of interaction form across the Double Diamond phases

The figures show the distribution of prompts in the different phases of the Double Diamond according to the two levels of coding. Figure 1 shows the second-level coding, which distinguishes between informational and generative requests, while Figure 2 shows the first-level coding, relating to the form of interaction with the tool.

As the two figures show, in the divergent phases, Discover and Develop, requests made to AI are almost exclusively informational, representing over 90% of requests in both phases. Furthermore, at least half of the interactions in these two phases are formulated as questions. This suggests that AI is used in an exploratory mode oriented toward information seeking.

In the convergent Define phase, generative requests prevail, accounting for 79% of the total, while over 50% of interactions take an imperative form. This indicates the use of AI in executive mode, where the tool is employed more directly to produce design outputs.

As for the Deliver phase, the results are not examined, since only 3 interactions out of 188 were recorded during this stage, a number too limited for interpretation.

3.2 Cross-check between request themes and interaction intentions

The results of the intention analysis were cross-referenced with those of the thematic analysis of student-AI interactions to verify whether requests classified as informational or generative were consistent with the actual content of the prompts. As Table 1 shows, informational requests are mainly associated with topics such as case studies or product examples. Generative requests, on the other hand, mainly concern the creation of personas and brief writing.

Table 1: Cross-check between thematic categories and request types in student–AI interactions

Theme	Informational	Generative
Case study	36	0
Product example	29	4
Technical support	31	1
Manufacturing techniques	29	0
Personas	3	20
Brief	0	10
Guidelines	3	5
Not classifiable	7	3

This cross-check supports consistency between the analysis of intentions and that of themes. Interactions that, based on the type of request and form of interaction, suggested the use of AI in exploratory mode are effectively oriented toward information searching, while those that suggested an executive mode are more often aimed at producing design outputs.

The Not classifiable category includes those prompts that could not be assigned to any of the thematic categories that emerged, because they were semantically difficult to interpret or not numerous enough to constitute a new category. For reasons of readability, the table shows only the main themes, excluding two less frequent categories, Moodboard and System-product scenario. The two unspecified requests have also not been included in the table and, if considered, would fall into the Not classifiable category.

3.3 Perceived usefulness and emotional experience of AI use

The results of the student–AI intention analysis were triangulated with questionnaire data to understand how different modes of AI use influenced students' experience, both in terms of perceived usefulness and emotional responses.

As Table 2 shows, the phases in which AI was used primarily in an exploratory mode, Discover and Develop, are those in which the tool was perceived as most useful. Perceived usefulness, however, decreases in the Define phase, where AI was used in executive mode. The sentiment score also follows a similar trend: it is positive in the divergent phases, neutral and negative in the convergent phases. Overall, these results suggest that the exploratory mode is associated with greater perceived usefulness and a more positive emotional experience.

Table 2: Perceived usefulness of AI and sentiment score across the Double Diamond phases

Phases	Mean perceived usefulness	Sentiment score
Discover	2.81	1
Define	2.44	0
Develop	2.86	4
Deliver	1.50	-1

Figure 3 shows the distribution, for each phase of the Double Diamond, of the emotional responses that emerged from the students' open-ended answers regarding their experience using AI.

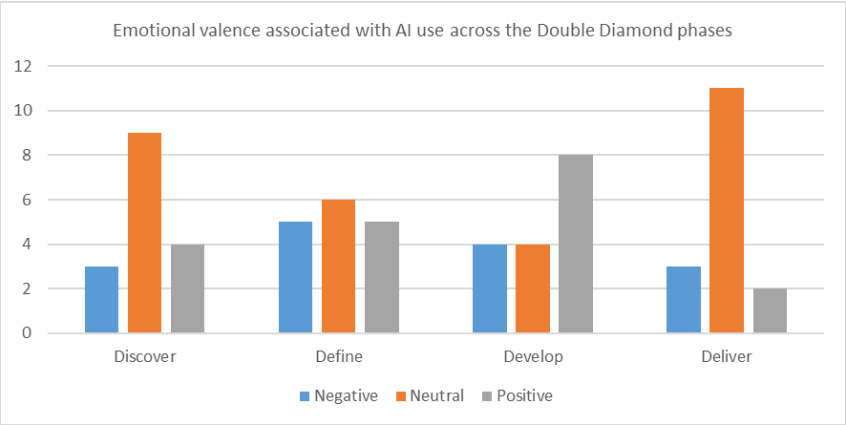


Figure 3: Distribution of emotional valence associated with AI use across the Double Diamond phases

The figure shows how positive, negative, and neutral responses are distributed across the various phases of the process. Responses classified as positive are primarily associated with the perception of having felt supported by the AI during the activity. Responses classified as negative, on the other hand, are primarily linked to experiences of limitation or discouragement when using the tool. Finally, neutral responses include ambivalent situations, such as feeling supported but limited, or supported but discouraged.

More specifically, feelings of discouragement are mostly linked to difficulties in interacting with the tool, particularly when trying to get the AI to understand the desired objective. The perception of limitations, on the other hand, is mainly attributed to outputs that are considered inadequate or fail to meet students' expectations.

4 Discussion

This study interprets interactions between students and AI as a bidirectional relationship, in which humans and AI collaborate in the construction and transformation of knowledge. This interpretation aligns with the perspective that generative AI not only acts as a tool to improve task efficiency but also participates in the dynamics of knowledge construction (Chen et al., 2023; Cress & Kimmerle, 2023). From this perspective, the identification of two prevalent modes of AI use, exploratory and executive, suggests that the tool can support different knowledge processes.

During the divergent phases, AI is used in an exploratory mode, where students employ it to search for information; this appears to be associated with knowledge acquisition processes. In the convergent phases, AI appears to be used in an executive mode, with requests focused on generating design outputs. This second mode appears closer to a phase of knowledge application, in which students use both prior knowledge and

knowledge acquired in the exploratory mode to produce, evaluate, and refine the outputs generated by AI. This interpretation is consistent with the literature highlighting how AI can support knowledge-building processes in education, through practices oriented toward both the creation and application of knowledge (Chen et al., 2023; Tsoi & Strønen, 2024).

Figure 4 presents an interpretive model that links the identified modes of AI use to the dynamics of knowledge construction. The two modes identified are interpreted not as strictly separate uses, but as prevailing configurations of student–AI interaction within a dynamic process, in which AI can assume different functions depending on the cognitive demands of the phase.

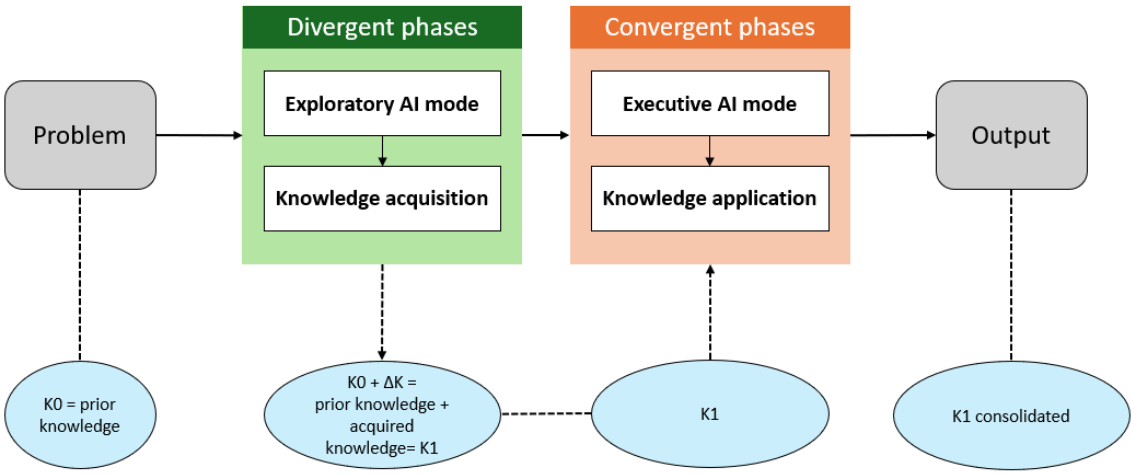


Figure 4: Interpretive model of student–AI interaction and knowledge building across the design process

Students begin with a challenge or problem with prior knowledge. During the divergent phases, they use AI in an exploratory mode to search for information, acquiring new knowledge that is combined with what they already know. In the convergent phases, AI is used in an executive mode to produce design outputs, which are evaluated and adapted based on the integration of prior knowledge and newly acquired knowledge, until the final result is achieved.

The model draws inspiration from the concept of absorptive capacity introduced by Cohen and Levinthal (1990), according to which learning depends on the ability to recognize the value of new knowledge, assimilate it, and apply it based on an existing knowledge base. Although this concept was formulated at the organizational level, it already assigned a central role to individuals. Subsequent developments on the topic then clarified this perspective at the individual level, leading to the concept of individual absorptive capacity. Sjödin et al. (2018) developed a model describing how individuals engage in the processes of absorbing new external knowledge. In this study, this perspective is reinterpreted and related to the ways in which students interact with generative AI systems within the design process.

This interpretation, however, should be treated with caution with regard to the convergent phases. In fact, even though the executive mode emerges quite clearly in the Define phase, it is not possible to extend this interpretation with the same certainty to the Deliver phase, because the number of interactions observed in the last phase was too limited to determine which mode of AI use actually prevails.

Furthermore, several limitations of the study should be considered. First, the small sample size of students limits the generalizability of the results. Second, the use of a single tool, KOBI, and its specific operational characteristics may have influenced both how the AI was used and the students' reported experience. Finally, the specific context of design and the adoption of the Double Diamond method may have encouraged modes of interaction with the AI that would not emerge in other disciplinary domains or in less design-oriented activities.

5 Conclusion

This study shows that the use of AI is not uniform but varies depending on the cognitive nature of the task. In particular, divergent activities are associated with an exploratory use of AI, in which students formulate prompts in interrogative form and the requests are oriented toward information-seeking. Convergent activities, on the other hand, are associated with an executive use of AI, in which prompts are primarily in the imperative form and requests are focused on generating design outputs. The results also show that the different modes of

AI use influence students' experience. Specifically, the exploratory mode was perceived as more useful and was associated with a positive sentiment, linked to a feeling of support during the activity.

This work contributes to a better understanding of how students interact with AI and how they experience that interaction. Furthermore, it proposes an interpretive model, inspired by the concept of absorptive capacity, that links the identified modes of AI use to the process of knowledge construction. The findings may provide useful guidance for the design of AI tools in education, which should be developed according to user-centered principles and taking into account the social, emotional, and cognitive aspects of the learning experience (Tlili et al., 2023; Muzumdar & Cheemalapati, 2025). At the same time, these findings are relevant because they provide insights for integrating AI more consciously and effectively into educational contexts, supporting students' learning processes (Lodge, 2023; Strzelecki, 2024; Muzumdar & Cheemalapati, 2025).

Future research could further examine how AI is used within the Double Diamond process, taking into account the limitations of this study. In addition, it would be useful to extend the research beyond the field of design, exploring other disciplinary contexts characterized by complex processes, in order to understand whether additional modes of interaction with AI might emerge. This understanding is fundamental to designing educational environments in which such tools do not just support task execution but contribute to learning processes.

6 Ethics declaration

Formal ethical clearance was not required for this study. The research was conducted in the context of an educational workshop, and the data were collected and analyzed in anonymized form. Participation in the questionnaires and in the use of interaction data for research purposes was voluntary.

7 AI declaration

The authors declare that ChatGPT was used to improve the readability and clarity of the text. All content has been reviewed and revised by the authors. The authors take full responsibility for the accuracy and originality of the work.

References

Ahmed, S. A. (2025). Learning, meaning, and teaching in the age of AI: Communication development checkpoints. *Frontiers in Education*, 10, 1720706.

Al-Zahrani, A. M., & Alasmari, T. M. (2024). Exploring the impact of artificial intelligence on higher education: The dynamics of ethical, social, and educational implications. *Humanities & Social Sciences Communications*, 11(1), 912.

Batista, J., Mesquita, A., & Carnaz, G. (2024). Generative AI and Higher Education: Trends, Challenges, and Future Directions from a Systematic Literature Review. *Information*, 15(11), 676.

Chen, B., Zhu, X., & Díaz Del Castillo, F. H. (2023). Integrating generative AI in knowledge building. *Computers and Education: Artificial Intelligence*, 5, 100184.

Cohen, W. M., & Levinthal, D. A. (1990). Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly*, 35(1), 128-152.

Cress, U., & Kimmerle, J. (2023). Co-constructing knowledge with generative AI tools: Reflections from a CSCL perspective. *International Journal of Computer-Supported Collaborative Learning*, 18(4), 607–614.

Design Council. (2005). Eleven lessons: managing design in eleven global brands. The design process.

Hagos, H. G., Lesjak, D., Flogie, A., & Babu, R. (2025). Artificial Intelligence in Knowledge Management for Higher Education: Transformative Impact, Challenges, and Future Directions Post-COVID-19. *Journal of Information Systems Engineering and Management*, 10(3s), 344-359.

Holmes, W., & Tuomi, I. (2022). State of the art and practice in AI in education. *European Journal of Education*, 57(4), 542-570.

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., Stadler, M., Weller, J., Kuhn, J. & Kasneci, G. (2023). ChatGPT for Good? On Opportunities and Challenges of Large Language Models for Education. *Learning and Individual Differences*, 103, 102274.

Lodge, J. M., Thompson, K., & Corrin, L. (2023). Mapping out a research agenda for generative artificial intelligence in tertiary education. *Australasian Journal of Educational Technology*, 39(1), 1-8.

Muzumdar, P., & Cheemalapati, S. (2025). Exploring Student Interactions with AI-Powered Learning Tools: A Qualitative Study Connecting Interaction Patterns to Educational Learning Theories. *Journal of Education and Training*, 13(2), 24.

Pulungan, N. K., & Nora, Y. E. (2025). Students' Experiences and Meaning-Making in AI-Based Adaptive Learning in Higher Education. *Journal of Educational Innovation and Research*, 1(10), 424-434.

Ripa di Meana, F., Giretti, A., Guidi, A., Vaccarini, M., & Zambelli, M. (2024). KOBI: A Knowledge Ecosystem for Research and Education. In A. Guidazzoli & M. C. Liguori (Eds.), *AI, Cultural Heritage, and Art. Between Research and Creativity*. (pp. 69-84).

Sahar, R., & Munawaroh, M. (2025). Artificial intelligence in higher education with bibliometric and content analysis for future research agenda. *Discover Sustainability*, 6(1), 401-420.

Schön, D. A. (1983). *The reflective practitioner: How professionals think in action*. Basic Books.

Sjödin, D. R., Frishammar, J., & Thorgren, S. (2019). How Individuals Engage in the Absorption of New External Knowledge: A Process Model of Individual Absorptive Capacity. *Journal of Product Innovation Management*, 36(3), 356-380.

Strzelecki, A. (2024). Students' Acceptance of ChatGPT in Higher Education: An Extended Unified Theory of Acceptance and Use of Technology. *Innovative Higher Education*, 49(2), 223–245.

Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R. & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments*, 10(1), 15

Tsoi, J. C., & Strønen, F. (2024). Integration of Conversational AI Capabilities in Knowledge Management Processes for Higher Education. Proceedings of the 25th European Conference on Knowledge Management (ECKM 2024).

Yusuf, A., Pervin, N., & Román-González, M. (2024). Generative AI and the future of higher education: a threat to academic integrity or reformation? Evidence from multicultural perspectives. *International Journal of Educational Technology in Higher Education*, 21, 21.

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? *International Journal of Educational Technology in Higher Education*, 16, 39.

Building Trust in Hybrid Teams: The Role of AI-Assisted Storytelling

Suncana Tuksar, Vanja Bevanda
Juraj Dobrila University of Pula, Faculty of Economics and Tourism "Dr. Mijo Mirkovic", Pula, Croatia

Abstract

The aim of this study is to develop a theoretically integrated conceptual framework for trust formation in AI- hybrid teams. Drawing from Media Synchronicity Theory (MST), Computer-Supported Cooperative Work (CSCW), hybrid intelligence and knowledge management literature, the study reconceptualizes trust as a communicative co-construct, which emerges from socio-technical interactions. Hybrid teams increasingly rely on the AI for collaboration, yet little is known about how *AI-mediated narratives* shape trust and the knowledge co-creation. This gap is addressed by examining how AI-assisted storytelling communicates and supports sensemaking in shaping both relational and knowledge-based outcomes. The objectives are: (1) to conceptualize AI-assisted storytelling as a shared collaborative artifact that supports knowledge co-creation; (2) to examine the mechanisms through which trust develops, focusing on narrative construction, perceived social proximity and AI transparency; (3) to specify the role of digital literacy in trust formation when engaging with AI-mediated narratives; (4) to integrate MST and CSCW perspectives into a unified framework of those communication processes; (5) to propose a testable conceptual model (H1–H5) that advances theoretical understanding of trust in hybrid teams. Methodologically, the study integrates the computational bibliometric analysis and structured theoretical synthesis in a two-phase conceptual review. Phase 1 combines AI-assisted literature mapping with the semantic clustering to identify relevant thematic clusters and gaps. Phase 2 uses a theory-driven deductive synthesis to identify mechanisms and create the conceptual model and associated hypotheses. The findings indicate that hybrid team collaboration is most effective when AI-mediated structures are integrated with human-driven storytelling. AI contributes scalability and structural consistency, whereas human actors provide contextual nuance and foster relational trust. The results further suggest that AI-assisted storytelling enhances engagement, comprehension, relational cohesion, and trust. Finally, the study proposes a testable conceptual model (H1–H5) that offers theoretical guidance and practical implications for future research to further examine and extend these mechanisms.

Keywords: Hybrid teams; AI-assisted storytelling; Trust formation; Digital literacy; Knowledge management; Human-AI collaboration.

Paper type: an Academic Research Paper

1 Introduction

In hybrid work environments, where teams rely on AI for collaboration, trust is actively developed through digitally mediated interactions rather than stemming from physical proximity. Perceived social proximity, digital literacy and transparent AI systems are essential for the engagement, collaborative knowledge construction and coordination. In this context, trust serves as a critical foundation for effective knowledge sharing, while AI-assisted storytelling serves as a critical mechanism for sensemaking, structuring communication and reinforcing relational cohesion. Nevertheless, little is known about how AI-mediated narratives develop, as well as about the socio-technical interplay between AI and human capabilities. While research on AI-mediated collaboration mainly focuses on communication efficiency, performance and technology adoption, trust as a dynamic, co-constructed phenomenon has received limited attention. Similarly, although storytelling and AI tools are individually recognized for their value, few studies integrate them to examine how AI-assisted narratives influence relational and knowledge outcomes. These gaps call for a socio-technical framework that systematically links narrative practices, trust formation, AI transparency and digital literacy in hybrid teams. This study develops a theoretically integrated framework to explain how AI-assisted storytelling shapes trust formation in hybrid teams. It advances understanding of AI-augmented collaboration by conceptualizing storytelling as a shared collaborative artifact and examining how trust emerges through narrative construction, perceived social proximity and AI transparency. It further considers the moderating role of digital literacy, integrates insights from Media Synchronicity Theory and CSCW and proposes a testable conceptual model to guide both research and practice. The methodology adopts a theory-driven structured literature review comprising two complementary phases. In Phase 1, AI-assisted exploratory mapping with semantic clustering is used to chart the literature on hybrid teams, trust, storytelling and AI-mediated collaboration. This process identified dominant thematic clusters and ensured a systematically selected corpus by combining automated clustering with manual theoretical validation. In Phase 2, a deductive, theory-informed synthesis was conducted to extract core mechanisms, operationalize constructs and develop a set of hypotheses. The integration of AI-

assisted bibliometric mapping and structured conceptual analysis reduced selection bias and ensured transparent theory development. As a result, the study proposes a testable conceptual model (H1–H5) to explain trust formation and collaborative knowledge co-creation in hybrid teams, offering a theoretically grounded and practically applicable framework for advancing AI-augmented hybrid collaboration.

2 Literature Review and Theoretical Foundations

2.1 Hybrid teams, communication, and trust

Hybrid teams, which combine co-located and remote members, introduce structural and communicative complexities that shape trust formation (Bennett et al., 2023). Because hybrid work is highly context-specific, sustaining effective practices requires strategies that integrate organizational culture, technological infrastructure and employee experience (Wiatr and Skowron-Mielnik, 2024). Trust is particularly critical in such settings, where limited face-to-face interaction may constrain cooperation and coordination (Hacker et al., 2019). Further, prior research demonstrates that knowledge-sharing strengthens trust and collaboration among virtual team members, with collaboration mediating the relationship between trust and team effectiveness (Alsharo et al., 2017). Perceived social proximity reduces spatial and psychological distance, thereby reinforcing relational cohesion and supporting effective knowledge exchange in hybrid teams (Bennett et al., 2023).

2.2 AI-assisted collaboration and hybrid intelligence

The AI integration fosters hybrid intelligence by combining human and machine capabilities to enhance decision-making and knowledge co-creation (Dellermann et al., 2019). Transparency and explainability are critical for trust in human-AI collaboration, as uninterpretable systems can undermine confidence. In hybrid teams, adaptive interaction and trust-supportive AI design are necessary to balance predictability with autonomy (Klein et al., 2021). Trust, reinforced by organizational support, strengthens engagement and performance in hybrid contexts (Wahyutomo et al., 2024). Overall, effective AI-assisted collaboration reflects a socio-technical integration of system capabilities and relational dynamics that shape team outcomes.

2.3 Storytelling as knowledge co-construction

Storytelling enhances understanding, alignment, and behavioural change within teams (Trabucchi et al., 2022), while also enabling coordinated action and strengthening interpersonal connections – an especially critical function in hybrid contexts. Prior research shows that stories facilitate mutual understanding and trust by providing a shared frame of reference (Brown et al., 2019). Extending this perspective, AI-assisted storytelling supports the structuring, and integration of narratives (Javadi et al., 2024), as well as bridging temporal, spatial, and cognitive gaps in collaborative settings (Klein et al., 2021).

2.4 The moderating role of digital literacy

Digital literacy plays a critical moderating role in engagement with AI and collaborative technologies, strengthening trust and relational cohesion. Empirical evidence indicates that technical competence strongly predicts collaborative technology use ($\beta = .816$, $p < .001$) and perceived social proximity ($\beta = .311$, $p < .001$) (Bennett et al., 2023). Core competencies include information processing, communication, and critical thinking (European Commission, 2013, 2022). Recent research further suggests that digital literacy enables effective human–AI co-creation by enhancing interpretability, engagement, and trust within socio-technical interactions (Javadi, 2024; Tony et al., 2024; Chen et al., 2024; Kholina and Frolova, 2024).

2.5 Trust as a communicatively constructed phenomenon

Trust is constructed through ongoing interaction rather than existing a priori (Hao et al., 2022; Hacker et al., 2019). Prior research establishes a strong positive relationship between trust and knowledge sharing (Staples & Webster, 2008). In AI-mediated hybrid contexts, trust develops through shared narrative construction (Trabucchi et al., 2022), perceived social proximity (Bennett et al., 2023), AI transparency (Vodiskar and Ruiner, 2025), and digital literacy (Bennett et al., 2023; European Commission, 2022; Kholina and Frolova, 2024).

2.6 Theoretical integration

Although some foundational studies fall outside the scope of the current review, they provide essential theoretical grounding for the present framework. MST links communication effectiveness to task-media alignment by distinguishing between conveyance and convergence processes. AI-assisted storytelling supports both processes by structuring information and facilitating shared understanding. Complementing this perspective, CSCW emphasizes the role of technological mediation in enabling collaboration (Klein et al., 2021), with AI-generated narratives bridging spatial and temporal gaps (Trabucchi et al., 2022). The integration of human and AI capabilities thus gives rise to HI, in which co-created knowledge artifacts enhance decision-making, creativity, and team problem-solving (Dellermann et al., 2019).

2.7 Research gap

Despite growing research on hybrid collaboration and AI, trust as a communicatively constructed phenomenon in AI-mediated storytelling remains underexplored (Alsharo et al., 2017; Hincapie and Costa, 2024). Much of the existing literature emphasizes efficiency and technology adoption, often overlooking relational and interpretive processes that underpin effective collaboration. This gap emphasizes the importance of understanding trust in the development of a framework that links storytelling, digital literacy and socio-technical dynamics "to support hybrid intelligence and collaborative knowledge creation" (Trabucchi et al., 2022).

3 Methodology: Theory-Driven Structured Literature Review

3.1. Research design

As shown in Table 1, the study employed a hybrid-structured conceptual review to ensure thematic relevance and contemporary applicability, comprising two complementary phases. Phase 1 used AI-assisted literature mapping with semantic clustering to identify key research clusters in hybrid teams, trust, and storytelling. Phase 2 applied a theory-driven conceptual review to analyse retained literature and develop hypotheses. Together, this approach integrates AI-assisted bibliometrics with deductive synthesis to support the conceptual model.

Table 1: Literature distinction model for screening criteria

Phase	Purpose / Role	Relevant Paper Section
Phase 1: AI-assisted literature mapping	Exploratory and cartographic; Identifies thematic clusters; Charts the systematic and unbiased coverage of literature	Methodological; Demonstrates literature selection and corpus mapping
Phase 2: Theory-driven structured conceptual review	Deductive and synthesis-oriented; Extracts mechanisms; Develops hypotheses; Builds the conceptual model	Theoretical Background; Provides conceptual and theoretical foundations for the conceptual model.

3.2 Phase 1: AI-assisted literature mapping and thematic clustering

Phase 1 examined the conceptual background of hybrid teams, trust, storytelling and AI-mediated collaboration to identify gaps and inform the theory-driven synthesis of Phase 2. An initial set of 128 publications was retrieved and semantic clustering was performed using SciSpace Agent to detect conceptual linkages across the corpus. After removing duplicates (n = 21) and non-relevant publications (n = 33), 74 studies were retained. Inclusion required explicit attention to trust development in hybrid teams, mechanisms linking communication quality to performance, or storytelling in relational and leadership contexts. Clustering was conducted in two stages. Automated semantic grouping first organized publications into preliminary clusters, which were then manually refined for theoretical coherence. This process identified four dominant thematic clusters. The first, trust-building mechanisms in hybrid teams, emphasizes trust as multidimensional and dynamic. It develops through stages – formation, facilitation, establishment, acceleration, and reinforcement – rather than as a static construct. Early-stage cognitive trust is particularly important in partially distributed teams as the evidence shows that interactions between trust types and team efficiency evolve over time, with virtuality reshaping, but not weakening, trust dynamics. The second cluster, communication quality and performance translation, links micro-level interactions to macro-level outcomes. High-quality communication, shared mental models and collaborative practices mediate the relationship between trust and project success. The third cluster, storytelling as a relational and leadership mechanism, positions narratives as socio-cultural tools for identity construction, belonging and leader credibility, which indirectly builds trust. Across clusters, key interconnections emerged:

communication quality mediates the influence of storytelling and trust on performance and virtuality shapes the salience of trust dimensions. Phase 1 generated a visual literature network and a clustered reference dataset documenting thematic clusters and conceptual relationships. These outputs enhance transparency, reduce selection bias, and guided the Phase 2 deductive synthesis on AI-assisted storytelling, transparency, digital literacy, and hybrid intelligence.

Figure 1 illustrates the Phase 1 research design, depicting the systematic search, filtering, semantic clustering, and manual validation that produced four thematic clusters: trust dynamics, communication and performance mechanisms, storytelling in relational leadership, and socio-technical mediation.

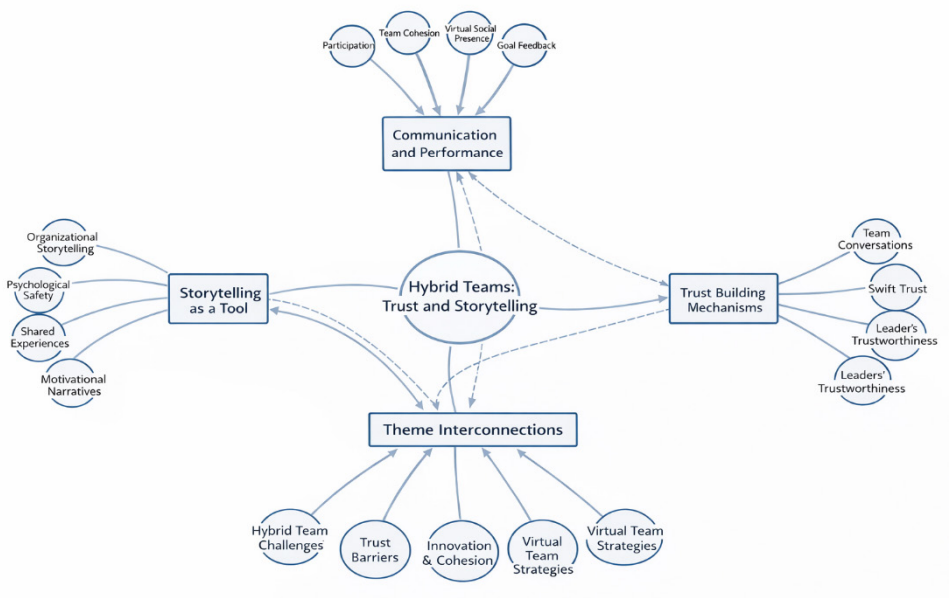


Figure 1: AI-assisted literature mapping

3.3 Phase 2: conceptual synthesis and model development

Phase 2 applies a theory-driven structured review to synthesize studies on AI-assisted storytelling, digital literacy and trust formation in hybrid teams. The review integrates MST, CSCW, HI and KM research. Foundational studies pre-2018 provide theoretical grounding, while empirical analysis focuses on 2018–2025, reflecting the acceleration of hybrid work post-COVID-19, adoption of applied AI and emergence of generative AI and human–AI collaboration systems. Phase 2 identifies mechanisms linking AI-assisted storytelling to trust, operationalizes constructs, formulates hypotheses (H1–H5), and ensures theoretical alignment across retained literature. Structured searches were conducted in Google Scholar, Scopus, and Web of Science (Core Collection) for their indexing quality and interdisciplinary coverage. A Boolean query combined thematic clusters to ensure relevance: ("hybrid teams" or "hybrid work" or "virtual teams") and ("AI-assisted collaboration" or "artificial intelligence" or "human–AI interaction") and ("storytelling" or "narrative") and ("trust" or "social proximity") and ("digital literacy" or "digital competence"). Searches were limited to titles, abstracts, and keywords, and included only peer-reviewed journal articles published over the last decade, with a particular focus on 2018–2025). Selected studies covered a range of collaborative, socio-technical, communicative and relational perspectives, ensuring methodological and contextual diversity.

Table 2 summarizes the Phase 2 literature by research approach, context and core mechanisms, illustrating methodological plurality and diversity and supporting mechanism extraction for the conceptual model.

Table 2: Literature base (Phase 2)

Research Approach	Context / Domain	Core Conceptual Mechanisms Identified
Quantitative surveys	Virtual and hybrid teams	Trust, knowledge sharing, social proximity
Experimental studies	Human-AI collaboration; AI storytelling; learning	AI transparency, shared narrative construction, engagement
Qualitative case studies	Organizational change; hybrid work design	Narrative alignment, relational cohesion, hybrid collaboration
Conceptual reviews	Human-AI teaming; virtual collaboration	Hybrid intelligence, socio-technical interaction, trust dynamics

Digital competence frameworks	Work; education; society	Digital literacy as enabling competence
Conceptual-experimental AI research	AI-mediated co-creation	Narrative curation, interpretive skills, collaborative knowledge creation

3.4 Analytical procedure

Phase 2 applied a theory-driven review with optimized Boolean searches to extract mechanisms and formulate hypotheses. The corpus comprised diverse methodologies, including conceptual analyses, surveys, case studies, experiments and structured reviews. This ensured comprehensive theoretical and empirical insights to support thematic synthesis, coding and model development. Cross-study comparison revealed underexplored intersections among communication, knowledge management and human-AI collaboration, indicating that trust is rarely examined as a dynamic, AI-mediated socio-technical process. To address this, a three-step integrative procedure was applied: thematic synthesis, conceptual coding, and theory mapping. The focus was on communication processes, digital literacy as a moderator and trust as a co-constructed dimension. This procedure directly guided the formulation of the hypotheses (H1-H5) and supported the development of the integrated conceptual model.

Figure 2 shows the systematic progression from exploratory mapping to focused synthesis. Phase 1 employed AI-assisted bibliometric mapping to explore the broader literature landscape and identify four thematic clusters.

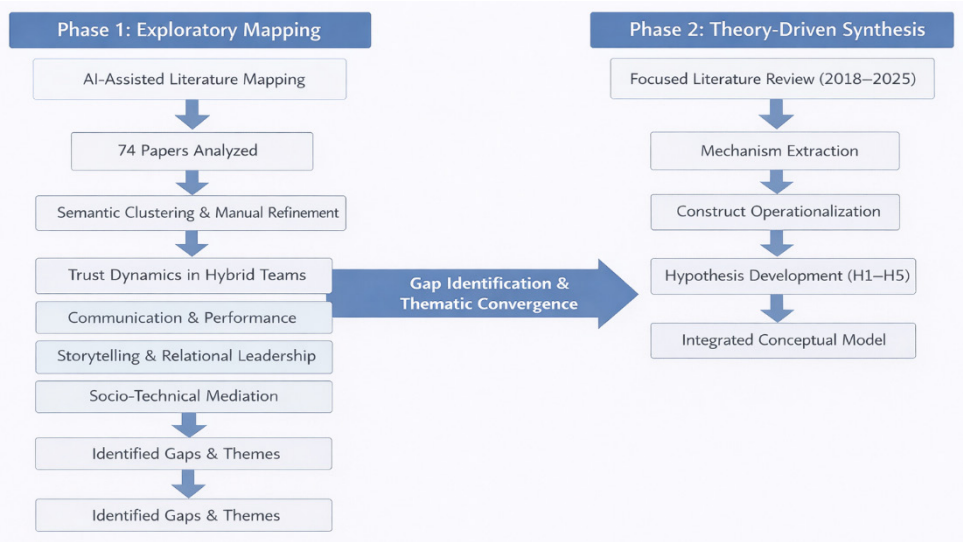


Figure 2: The two-phase research design for conceptual model development

4 Results: Hypotheses and the Conceptual Model

Following the structured literature synthesis, the results are formalized as theoretically grounded hypotheses and an integrated conceptual model.

4.1 Hypotheses

- ***H1:*** *AI-assisted storytelling functions as a shared collaborative artifact that facilitates collective sensemaking and positively influences trust in hybrid teams.*
- ***H2:*** *Perceived social proximity among hybrid team members is positively associated with trust in AI-mediated collaborative storytelling contexts.*
- ***H3:*** *The transparency and explainability of AI-assisted storytelling processes positively influence trust in hybrid team collaboration.*
- ***H4:*** *Digital literacy moderates the relationship between AI-assisted storytelling and trust, such that higher levels of digital literacy strengthen this relationship.*
- ***H5:*** *Hybrid team configurations characterized by higher degrees of digital mediation strengthen the positive effect of AI-assisted storytelling on trust formation compared to less digitally mediated contexts.*

4.2 Conceptual model

Based on the insights of the researchers in this study, the model frames trust as a co-constructed outcome of socio-technical interactions rather than a static antecedent or consequence. AI-assisted storytelling serves as a shared collaborative artifact that structures communication and supports collective sensemaking. Trust emerges through three mechanisms: shared narrative construction, which enables knowledge co-creation, perceived social proximity, which reduces spatial and psychological distance, and AI transparency and explainability, which reinforcing confidence in human-AI interaction. Digital literacy is theorized as a moderating capability, which assists in the engagement with AI-assisted storytelling and sustaining trust development. Hybrid team configuration further conditions these relationships, with greater digital mediation amplifying the impact of AI-assisted storytelling on trust. Overall, such framework integrates communication processes, technological affordances and human competencies into a coherent socio-technical structure, formalized in five analytically grounded and empirically testable hypotheses (H1–H5).

Figure 3 presents the conceptual model of trust formation in AI-augmented hybrid teams.

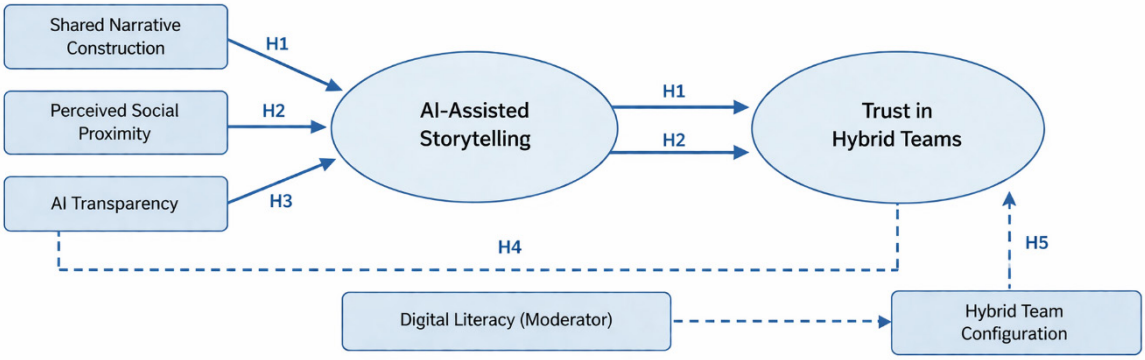


Figure 3: Conceptual model of trust formation in AI-augmented hybrid teams.

4.3 Conceptual mechanisms and hypotheses mapping

Table 3 summarizes how the literature review supports the mechanisms of the proposed model. Trust in hybrid teams emerges dynamically through socio-technical interactions, with human contributions, social proximity and AI transparency, while digital literacy enables the knowledge co-creation. The structure of a hybrid team further shapes these effects, as digitally mediated arrangements amplify the influence of AI-assisted storytelling. Together, these insights provide the empirical and conceptual foundation for hypotheses H1–H5.

Table 3: Hypotheses mapping

Hypothesis	Concept – Mechanism	Key Findings – Contributions
H1	Shared narrative construction	AI-assisted storytelling enables collective sensemaking and knowledge co-creation; Human input adds contextual nuance and adaptive reasoning
H2	Social proximity	Perceived closeness fosters engagement, relational cohesion and trust in hybrid teams
H3	AI transparency and explainability	Transparent AI outputs enhance predictability and confidence, supporting effective human-AI collaboration
H4	Digital literacy	Higher digital competence strengthens engagement with AI narratives, improves interpretation and sustains trust and relational cohesion
H5	Hybrid team configuration	Flexible, distributed team arrangements amplify AI-assisted storytelling effects, supporting trust, collaboration and knowledge sharing

4.4 Literature retention by conceptual constructs

The retained literature provides strong empirical and conceptual support for the following model constructs: shared narrative construction, social proximity, AI transparency, AI-assisted storytelling, digital literacy, and hybrid team configuration. Together, these constructs shape trust, engagement, and knowledge co-creation in AI-augmented hybrid teams and support the five proposed hypotheses (H1–H5).

4.5 Critical synthesis and gap analysis

Table 4 summarizes key mechanisms supporting AI-assisted hybrid collaboration and highlights their main contributions and critical gaps. Most studies examine mechanisms in isolation, leaving AI-assisted storytelling, digital literacy and team-trust interactions underexplored. However, this study addresses these gaps with a unified socio-technical framework for trust formation in AI-augmented hybrid teams.

Table 4: Synthesis and gap analysis

Mechanism	Strengths	Gaps / Limitations
Shared narrative construction	Supports knowledge co-creation, alignment and trust	Few studies on AI-assisted storytelling in hybrid teams
Perceived social proximity	Reduces distance; Fosters relational cohesion	Limited research on AI-mediated influence on social proximity
AI transparency and explainability	Enhances trust and effective human-AI collaboration	Lack of empirical quantification of transparency effects in hybrid teams
AI-assisted storytelling	Demonstrated as co-creative tool; Supports learning	Scarce longitudinal or large-scale studies; Novelty underexplored
Digital literacy	Strengthens engagement, interpretation and trust	Mostly self-reported; Limited intervention-based evidence
Trust in hybrid teams	Central to collaboration and knowledge sharing	Limited focus on trust in AI-mediated narrative practices
Hybrid team configuration	Shapes relational and communicative dynamics	Few studies link team structure with AI-assisted collaboration and trust

5 Discussion, Implications and Future Research

Based on the critical synthesis, AI-assisted storytelling is positioned as a central socio-technical mechanism driving trust in hybrid teams. Table 5 summarizes the key findings, theoretical contributions, and practical implications.

Table 5: Key findings, implications and mechanisms

Hypothesis – Aspect	Key Findings	Implications
H1: AI-assisted narratives	Facilitates co-construction of meaning, alignment and relational cohesion	Supports shared sensemaking, knowledge co-creation and team cohesion
H2: Perceived social proximity	AI-mediated narratives reduce spatial and psychological distance; Foster engagement	Enhances trust, relational cohesion and psychological safety in hybrid collaboration
H3: AI transparency and explainability	Transparent AI outputs increase predictability and trust	Highlights need for user-centered and more explainable AI systems
H4: Digital literacy	Higher digital competence improves engagement with AI outputs, narrative integration and trust	Justifies investment in digital skills to enable effective human-AI collaboration
H5: Hybrid team configurations	AI-assisted co-creation allows distributed teams to achieve outcomes comparable to co-located teams	Encourages strategic hybrid team design and AI-supported workflows
Overall socio-technical mechanism	Trust emerges dynamically from narratives, proximity, transparency and digital literacy (moderator)	Demonstrates the integrated socio-technical nature of AI-assisted hybrid collaboration

5.1 Key findings

AI-assisted narratives act as shared collaborative artifacts, facilitating alignment, mutual understanding and knowledge co-creation in hybrid teams. Social proximity and transparent AI outputs strengthen relational cohesion and trust, while digital literacy enables effective interpretation and integration of AI-generated narratives. Hybrid team configurations show that distributed teams can achieve outcomes comparable to co-located teams when AI-assisted storytelling is applied. Together, these mechanisms demonstrate that trust in hybrid teams is a dynamic, co-constructed socio-technical phenomenon.

5.2 Theoretical contributions

This study reconceptualizes trust in hybrid teams as a communicatively co-constructed phenomenon and positions AI-assisted storytelling as a knowledge artifact that fosters alignment and relational cohesion. It highlights digital literacy as a key moderating capability connecting human competencies to AI-mediated collaboration. It further demonstrates the integrated socio-technical nature of trust formation, where narrative artifacts, relational practices and technological affordances together generate knowledge and engagement outcomes.

5.3 Practical implications

The findings suggest that AI systems should prioritize transparency, explainability and user-cantered interaction. AI-mediated narratives can strengthen social proximity, engagement and relational cohesion in distributed teams, while digital literacy programs enhance interpretation of AI outputs and support collaborative knowledge co-creation. Structuring hybrid teams to leverage AI-assisted storytelling can thus improve both relational and knowledge-based performance.

5.4 Future research directions

Future research should empirically test the proposed model using surveys, experiments and field studies, including organizational, cultural and industry-specific moderators of trust formation. Longitudinal studies could illuminate the evolving dynamics of trust in AI-mediated hybrid teams, while further investigation into AI functionalities and storytelling interventions could optimize engagement, knowledge co-creation and socio-technical collaboration.

6 Conclusion

This study advances understanding of trust formation in AI-augmented hybrid teams by positioning AI-assisted storytelling as a central socio-technical mechanism that shapes collective sensemaking and co-created knowledge. Trust is reconceptualized as a dynamic, communicatively co-constructed outcome, which emerges from the interplay of narrative practices, relational interactions and technological affordances. Hereby, AI functions as an active collaborative partner, providing structure, interpretability and alignment, while human contributions deliver contextual nuance and relational trust. Digital literacy serves as a moderator which enforces the engagement with AI, as well as influences the integration of narratives and the development of trust. The conceptual framework highlights AI-assisted storytelling as a key facilitator of knowledge co-creation and identifies digital literacy as a critical competence in hybrid collaboration. Practically, it guides organizations in designing AI narrative practices and structuring hybrid teams to enhance relational cohesion and collaborative outcomes. Although conceptual and literature-based, the study lays a foundation for empirical research, including longitudinal, cross-cultural and tool-specific investigations, offering a socio-technical lens to understand and enhance trust in AI-augmented hybrid teams.

7 AI Declaration

ChatGPT Pro (OpenAI) was used to support the generation of Figures 1-3, which were subsequently reviewed and edited by the authors. The AI-generated material was used as a drafting aid for visualization and not as an autonomous source of scientific content. SciSpace Agent (TL;DR) was used to cluster and summarise an initial set of 128 publications during the literature review process. AI tools were employed solely as assistive instruments for visualization support and literature organization. No AI tool was used to generate final scientific claims, results or conclusions presented in the manuscript. All content has been reviewed, validated and approved by the authors.

8 Ethics Declaration

This research did not involve human participants, personal data, or animal subjects; therefore, no ethical approval from an institutional review board was required. All work complies with general ethical standards for scholarly research and publication.

References

Alsharo, M., Gregg, D. and Ramirez, R. (2017) "Virtual team effectiveness: The role of knowledge sharing and trust", *Information and Management*, Vol. 54, pp. 479-490.

Bennett, A.A., Campion, E.D., Keeler, K.R. and Keener, S.K. (2023) "Digital literacy, the use of collaborative technologies, and perceived social proximity in a hybrid work environment: Technology as a social binder", *Computers in Human Behavior Reports*, Vol. 9, 100247.

Chen, Q., Shuai, W., Zhang, J., Sun, Z. and Cao, N. (2024) "Beyond numbers: Creating analogies to enhance data comprehension and communication with generative AI",
[online], ACM Digital Library, https://dl.acm.org/conference/chi.

Dellermann, D., Ebel, P., Söllner, M. and Leimeister, J.M. (2019) "Hybrid Intelligence", *Business and Information Systems Engineering*, Vol. 61, No. 5, pp. 637-643.

Deschenes, A.A. (2024) "Digital literacy, collaborative technologies, and perceived social proximity in a hybrid work environment", *Digital Business*, Vol. 4, No. 1.

European Commission, Joint Research Centre (2013) *DIGCOMP 2.0: The Digital Competence Framework for Citizens*, Publications Office of the European Union, Luxembourg,
[online], JRC Publications Repository, https://publications.jrc.ec.europa.eu/repository/home.

European Commission, Joint Research Centre (2022) *The Digital Competence Framework for Citizens (DigComp)*, Publications Office of the European Union, Luxembourg,
[online], JRC Publications Repository, https://publications.jrc.ec.europa.eu/repository/home.

Hacker, J.V., Johnson, M., Saunders, C. and Thayer, A.L. (2019) "Trust in Virtual Teams: A Multidisciplinary Review and Integration", *Australasian Journal of Information Systems*, Vol. 23.

Hao, Q., Zhang, B., Shi, Y. and Yang, Q. (2022) "How does trust in coworkers foster knowledge sharing in virtual teams? A multilevel moderated mediation model of psychological safety, team virtuality, and self-efficacy", *Frontiers in Psychology*, Vol. 13, 899142.

Hincapie, M.X. and Costa, P. (2024) "Fostering hybrid team performance through inclusive leadership strategies", *Organizational Dynamics*,
[online], Science Direct, https://www.sciencedirect.com/science/article/abs/pii/S0090261624000457.

Javadi, V.S., Trippas, J.R., Lal, Y.K. and Flek, L. (2024) "Can stories help LLMs reason? Curating information space through narrative",
[online], Open Review, https://openreview.net/pdf/27f016297c6591899fa65ff4e34b5be966e800cc.pdf.

Klein, G., Wolf, R., Dellermann, D. and Söllner, M. (2021) "Human-AI Teaming: Design Principles for Hybrid Intelligence Systems", *Journal of Business Research*, Vol. 127, pp. 195-207.

Kholina, O. and Frolova, E. (2024) "A network communicative culture for future teachers: Development of digital literacy and communicative competence", *Cogent Education*, Vol. 11, No. 1, 2363678.

Radensky, M. (2024) "Mixed-initiative methods for co-creation in scientific research",
[online], ACM Digital Library, https://doi.org/10.1145/3635636.3664627.

Staples, D.S. and Webster, J. (2008) "Exploring the effects of trust, task interdependence, and virtualness on knowledge sharing in teams", *Information Systems Journal*.

Tony, R., Bimal, N., Dinny, R., Baburaj, S. and Joykutty, A.M. (2024) "Memory Lane: Enhanced learning using AI-driven storytelling", *Proceedings of the 5th International Conference on Innovative Trends in Information Technology (ICITIIT)*, 15-16 March 2024,
[online], IEEEXplore Conference Proceedings, https://ieeexplore.ieee.org/document/10580203.

Trabucchi, D., Buganza, T., Bellis, P., Magnanini, S., Press, J., Verganti, R. and Zasa, F.P. (2022) "Story-making to nurture change: Creating a journey to make transformation happen", *Journal of Knowledge Management*, Vol. 26, No. 11, pp. 427-460.

Vodiskar, T. and Ruiner, C. (2025) "On human–AI collaboration and the meaningfulness of work", *Journal of Business Ethics*, (forthcoming).

Wahyutomo, D., Hakim, L. and Oktayani, D. (2024) "Driving Employee Performance in Hybrid Work Environments: The Role of Engagement, Support, Trust, and Respect", *Journal of Business Management and Economic Development*, Vol. 3, No. 01, pp. 120-130.

Wiatr, A. and Skowron-Mielnik, B. (2024) "The quest for a perfect hybrid working: A case study of Mars Wrigley Poland", *Human Systems Management*, Vol. 43, No. 6.

ChatGPT Acceptance among Higher Education Students: A Simplified Model

László Berényi
Institute of Management Science, University of Miskolc, Miskolc, Hungary

Abstract

Understanding and predicting the use of generative AI chatbots is an urgent task due to the rapid spread and rapid changes in technology. SEM models are popular in the field because they can account for the complex effects of multiple influencing factors. A serious problem is the data requirement of the analysis. Long surveys and large sample sizes are difficult to make feasible for the numerous targeted questions that arise in connection with the acceptance of generative AI use. The education system is significantly transformed by the emergence of AI. Strengthening its role in learning support and addressing academic misconduct requires actions grounded in readily accessible, up-to-date information. New framework models may support effective strategy building. The study aims to present a pilot framework model derived from existing models (Berényi et al., 2025) to measure the intention to use AI tailored to higher education students. The reflective measurement constructs are defined for PLS-SEM analysis. The limited number of items in the model enables rapid data collection by reducing the risk that respondents will lose interest while completing the survey. The pilot study for model validation used a sample of 281 students at the University of Miskolc (Hungary) and focused on ChatGPT use. SmartPLS4 was used for analysis. The pilot run confirmed acceptable reliability, with AVE between 0.626 and 0.814, HTMT <= 0.860, and VIF <= 1.974. Model fit is good (SRMR = 0.070). The model explained 55.9% of the total variance in intention to use and 47.2% of the total variance in usefulness. Path coefficients are significant, except for the impact of price-value. The results suggest that information quality impacts usefulness through usability; the direct impact is remarkably lower. The total impact of usability on intention to use is high (path coefficient is 0.629), partly through usefulness. Among the respondents, the impact of pricing on the intention to use was not significant, which may be due to free academic access to the chatbots or to the hoped-for profit exceeding the price. The results contribute to the body of knowledge on the acceptance of AI technologies among higher education students, but they can easily be extended to other purposes if it is assumed that intention to use depends on the quality of the information. The limited survey length increases the likelihood of completion, and the policymakers can work with more reliable information.

Keywords: technology acceptance, information quality, chatbot, PLS-SEM, higher education

Paper type: Academic Research Paper

1 Introduction

The introduction of generative artificial intelligence (genAI) chatbots led to remarkable changes in academic life and beyond. Although artificial intelligence has been incorporated much earlier into various services, planning assignments, and task-automation solutions, its popularity has increased with the spread of large language model (LLM)- based chatbots. The way of asking questions and receiving responses has become "human-like," making communication with the software easy. Since students realized that genAI can support writing homework, analyzing data, and solving many tasks with less effort than before, new challenges have emerged. Academic integrity recently includes AI-based cheating, or in a broader sense, the ethical use of genAI. Policymaking usually follows technological development. Effective policymaking requires understanding the complex motivations and habits of students. Second-generation multivariate statistical analysis offers reasonable opportunities in the field. Framework models such as the Technology Acceptance Model (TAM) (Davis, 1989; Venkatesh and Bala, 2008), the Unified Theory of Acceptance and Use of Technology (UTAUT) (Venkatesh et al., 2003; Venkatesh et al., 2012), or the Delone and McLean Information Systems Success Model (DM) (Delone and McLean, 1992; 2003) can be adopted to explore the factors influencing genAI use. However, the data collection survey is usually very long and detailed to ensure the consideration of several factors, and large samples are collected to ensure statistical validity. At the same time, the rapid development of technology makes results quickly outdated. Repeated surveys may lead to response bias, and students may lose interest in completing long questionnaires. The study aims to propose and validate a simplified framework suitable for repeated measurements in a rapidly evolving context. The model allows exploring the main influencing factors and comparing results over time, and across genAI versions. Using this model, the research goal is to measure the factors influencing ChatGPT use among higher education students. The results are based on a pilot study conducted with 281 business students at the University of Miskolc (Hungary).

The rest of the paper is organized into five main sections. Literature review highlights the appreciation of genAI in higher education. The materials and methods section presents the process of framework model development and the characteristics of the pilot sample. The results section gives an overview of the PLS-SEM modeling, followed by the discussion of the results. The conclusions include the main implications and limitations of the study.

2 GenAI in Higher Education

The spread of genAI chatbots meets the requirements imposed on universities to improve productivity. Publication pressure on teachers and researchers, as well as the need to complete students' mid-term tasks, may motivate people to use chatbot services to save time and effort. Hussein et al. (2025) emphasized that education is among the most frequently discussed themes in the context of genAI use. Considering that higher education prepares the next generation of engineers, medical doctors, economists, and philosophers, the spread of AI technology has a multiplicative impact on the economy and society.

The Student Generative AI Survey 2025 reported that 92% of UK undergraduate students used genAI, and the ratio has increased from 66% in 2024 (Freeman, 2025). The main reason for genAI covered saving time by summarizing content or generating ideas, as well as explaining information. Additionally, improving the quality of work through corrections was marked in the study.

Based on international comparisons, Annamalai et al. (2025) found a positive impact of genAI use on learning, including positive experiences with learning and the development of critical thinking. By offering styles explored by large language models, academic writing can be improved beyond grammar corrections (Ayoubi et al., 2024).

Although the opportunities are promising, cautious use must be considered due to mixed lessons (Farrokhnia et al., 2024; Zhou et al., 2024; Zábrátzky, 2025). The urge of institutional policies (Yang, 2024) is driven by the risks of falling academic work quality, writing skills, and critical thinking due to extensive reliance on genAI chatbots (Medina et al., 2024). At the same time, Jahani et al. (2024) emphasized the positive impact of genAI use on students' motivation, and Mohamed et al. (2025) highlighted its support for writing skills through grammatical and lexical corrections. Moreover, ethical issues must be considered when integrating genAI into education (Kumar et al., 2024). Kovari (2025) has proposed comprehensive strategic guidelines for reducing genAI risk. Since large language model-based chatbots can be considered emerging technologies, we cannot yet see their true impact despite numerous studies and opinions; more time and experience are needed. However, this does not mean that efforts to understand are pointless. Diverse opinions may be traced to local characteristics, but the wide variety of measurement models prevents a comprehensive comparison of results.

In the Hungarian context, Berényi and Csiszárik-Kocsir (2025) found generational differences in the use of genAI. Older generations are characterized as cautious users focused on professional tasks, while younger generations are confident users who have integrated genAI into their everyday activities. According to the scope, research, information synthesis, and translation are the main fields of use, similar to international surveys.

3 Materials and Methods

3.1 Framework model and hypothesis development

The research framework model (Figure 1) established intention to use ChatGPT as an exogenous construct, and information quality, usability, and usefulness as endogenous ones. Additionally, the impact of price-value was included. The model was designed for PLS-SEM analysis, following the instructions of Hair et al. (2022).

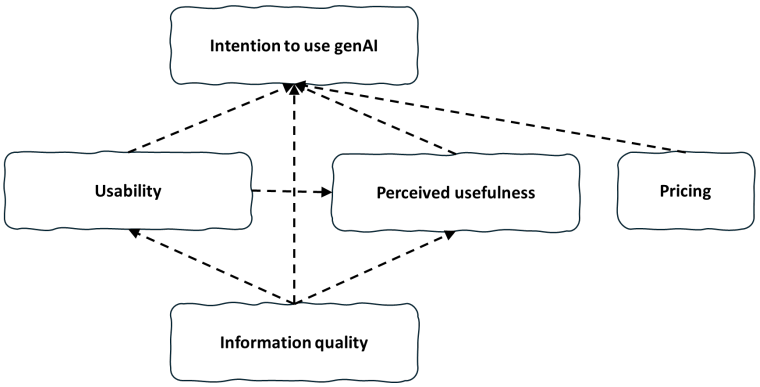


Figure 1: This is how a figure should appear

The latent variable constructs were defined as follows:

Information quality (IQ): the construct was adopted from the DM model and refers to the usefulness, accuracy, reliability, and comprehensibility of the information produced by an information system, as well as the extent to which it supports user decisions. (Delone and McLean, 2003).

Usability (US): is developed by the author, considering facilitating conditions in UTAUT2 (Venkatesh et al., 2012) and perceived ease of use in TAM (Venkatesh and Bala, 2008), referring to the perceived usability and enjoyment of the genAI chatbot.

Usefulness (PU): refers to the belief that using the genAI will contribute to performing tasks more efficiently.

Price-value (PV): is adapted from the UTAUT2 model (Venkatesh et al., 2012) and refers to the perceived return on investment (fees) of using genAI.

Intention to use (IU): is a general construct in similar models. The construct in the model refers to the intention of the respondents to use genAI for educational purposes and beyond in the future.

The research question was formulated as follows: How do information quality, usability, and perceived usefulness influence the intention to use ChatGPT? A methodological question is whether the provided framework model is applicable in the field.

There were six hypotheses formulated following the paths in the model:

- *H1: Information quality provided by ChatGPT affects its perceived usability.*
- *H2: Information quality provided by ChatGPT affects its perceived usefulness.*
- *H3: Perceived usability of ChatGPT affects perceived usefulness.*
- *H4: Information quality provided by ChatGPT affects the students' future intentions to use it.*
- *H5: Perceived usability of ChatGPT affects the students' future intentions to use it.*
- *H6: Perceived usefulness of ChatGPT affects the students' future intentions to use it.*
- *H7: Price-value of ChatGPT affects the students' future intentions to use it.*

The survey items selected from the previous analysis (Berényi et al., 2025) and their coding for the analysis section are summarized in Table 1.

Table 1**:** Survey items and coding

Construct	Item Code	Survey item
Information Quality	IQ1	ChatGPT offers reliable information
	IQ2	The information from ChatGPT is useful and easy to understand
	IQ3	Information from ChatGPT is in a form that is readily usable
Usability	UA1	I know how to find solutions for my problems about using ChatGPT features
	UA2	Using ChatGPT makes me satisfied in general
	UA3	I use ChatGPT for various solving tasks
	UA4	I spend a lot of time with ChatGPT
Usefulness	PU1	I think ChatGPT is useful for my studies
	PU2	ChatGPT can help improve the quality of my work
Price-value	PV1	ChatGPT is available at a reasonable price
	PV2	Pricing options of ChatGPT are flexible
Intention to use	IU1	I intend to continue using ChatGPT for my studies in the future
	IU2	I predict I will use ChatGPT beyond my studies

3.2 Data collection and sample

Data collection was conducted through a multipurpose online survey designed to assess ChatGPT use among higher education students in the fall semester of the 2024/2025 academic year. The questions were prepared in Google Forms using a 7-point scale. The study processes the responses of students at the University of Miskolc (Miskolc, Hungary). A total of 281 responses were available for analysis after filtering for academic misconduct, partially completed surveys, responses with very low standard deviation, and suspicious respondents. The representativeness of the sample was not checked. The sample consists of 73 female and 208 male students. 66.9% are aged 18-25 years, 18.5% are aged 26-35 years, and 14.6% are aged 36 years or older.

3.3 Method selection

Based on the guidance of Hair et al (2022), the PLS-SEM method was selected for the analysis. As the study aimed to develop a methodological pilot for a simplified theoretical approach to technology acceptance; the

sample size, the model structure, and the distribution of responses suggest using a PLS-SEM model. SmartPLS 4 software (Ringle et al., 2024) was used for the calculations. The bootstrapping procedure was performed with the bias-corrected and accelerated (BCa) method with 5000 subsamples. The significance level in the study was 0.05.

Based on the inverse-square-root method proposed by Kock and Hadaya (2018), the sample size is acceptable. The method suggests a minimum sample size n_{min} > (2.486(/p)2, where p stands for the minimum path coefficient in the PLS path model expected to be statistically significant. The study's sample size allows p = 0.148 as a minimum, which is lower than the value of 0.264 obtained during the run.

4 Results

4.1 Model validation

The descriptive statistics for model elements, their outer loadings, and the Variance Inflation Factor (VIF) values are summarized in Table 2. The outer loadings exceed the threshold of 0.708, except for IQ1, which equals 0.699. Since the value is close to acceptable and dropping the item would negatively impact model fit, it was retained in the model. The maximum VIF of 1.974 is less than 3, so there are no multicollinearity issues.

Table 2**:** Descriptive statistics of survey items, outer loadings, and VIF (based on SmartPLS output)

	Mean	Standard deviation	Excess kurtosis	Skewness	Outer loadings	VIF
IQ1	4.075	1.273	-0.185	-0.224	0.699	1.259
IQ2	5.235	1.345	-0.037	-0.620	0.858	1.628
IQ3	4.996	1.536	-0.443	-0.521	0.831	1.530
IU1	5.537	1.537	0.311	-0.984	0.912	1.717
IU2	5.609	1.527	0.605	-1.124	0.903	1.717
PV1	4.281	1.356	-0.127	-0.134	0.925	1.487
PV2	4.039	1.430	-0.255	-0.062	0.840	1.487
UA1	4.352	1.743	-0.824	-0.239	0.858	1.974
UA2	4.758	1.509	-0.385	-0.496	0.732	1.467
UA3	5.053	1.739	-0.629	-0.629	0.829	1.729
UA4	3.021	1.741	-0.668	0.573	0.739	1.506
PU1	5.413	1.488	0.036	-0.843	0.918	1.654
PU2	4.836	1.662	-0.636	-0.486	0.886	1.654

The reliability of the constructs is good (Table 3). Cronbach's alpha values are between 0.716 and 0.801, composite reliability indicators are higher than 0.7, and the average variance extracted (AVE) exceeds 0.5 in each case.

Table 3**:** Reliability analysis (based on SmartPLS output)

	Cronbach's alpha	Composite reliability (rho_a)	Composite reliability (rho_c)	AVE
Information quality	0.716	0.737	0.840	0.639
Intention to use	0.785	0.786	0.903	0.823
Price value	0.728	0.793	0.877	0.781
Usability	0.801	0.820	0.870	0.626
Usefulness	0.772	0.785	0.897	0.814

The model's discriminant validity was confirmed using the Fornell-Larcker criteria (Table 4) and the Heterotrait-Monotrait Ratio (HTMT) analysis (Table 5). The root squares of AVE indicators per construct are higher than the correlations with other constructs. According to HTMT, all values are below the threshold value of 0.85.

Table 4**:** Fornell-Larcker criteria (based on SmartPLS output)

	Information quality	Intention to use	Price value	Usability	Usefulness
Information quality	0.799				
Intention to use	0.398	0.907			
Price value	0.176	0.205	0.884		
Usability	0.458	0.685	0.228	0.791	
Usefulness	0.504	0.670	0.182	0.646	0.902

Table 5: Heterotrait-monotrait ratio (HTMT) analysis (based on SmartPLS output)

Relation	Heterotrait-monotrait ratio (HTMT)
Intention to use <-> Information quality	0.522
Price value <-> Information quality	0.239
Price value <-> Intention to use	0.265
Usability <-> Information quality	0.609
Usability <-> Intention to use	0.843
Usability <-> Price value	0.296
Usefulness <-> Information quality	0.672
Usefulness <-> Intention to use	0.849
Usefulness <-> Price value	0.240
Usefulness <-> Usability	0.810

The Standardized Root Mean Square Residual (SRMR) indicators suggest acceptable model fit ($SRMR_{saturated}$ = 0.700, $SRMR_{estimated}$ = 0.800). Other indicators suggest caution: NFI = 0.684, d_ULS = 0.448, d_G = 0.279, and Chi-square = 486.965 for the saturated model.

4.2 Model results

The path coefficients with bootstrapping are presented in Figure 2. Information quality (IQ) has a significant impact on Usability (UA) (ß = 0.458) and Usefulness (PU) (ß = 0.264), while its direct effect on the intention to use (IU) is negligible. Both UA (ß = 0.425) and PU (ß = 0.388) have significant effects on IU. The effect of price-value (PV) in IU (ß = 0.038) was not confirmed by the study.

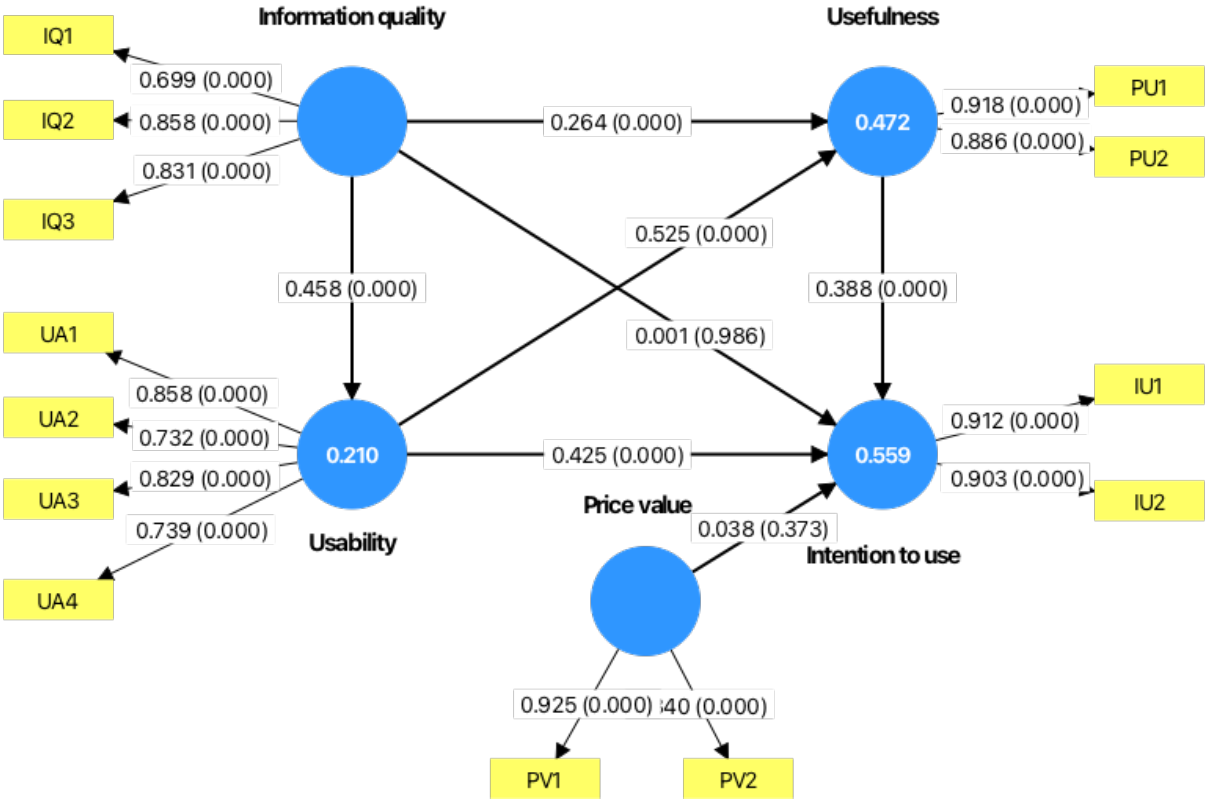


Figure 2: Path analysis results with significance levels (SmartPLS output)

The model explained 55.9% of the variance of the intention to use ChatGPT. IQ affects IU through UA and PU. The path coefficient between IQ and UA is higher than between IQ and PU, but the variance explained is the opposite. The path coefficient between UA and IU is higher than that between PU and IU (Table 6).

Table 6: Total effects and indirect effects (based on SmartPLS output)

	Total effect	p values	Total indirect effects	p values
IQ -> IU	0.391	0.000	0.390	0.000
IQ -> UA	0.458	0.000		
IQ -> PU	0.504	0.000	0.240	0.000
PV -> IU	0.038	0.373		
UA -> IU	0.629	0.000		
UA -> PU	0.525	0.000	0.204	0.000
PU -> IU	0.388	0.000		

Table 7: Path coefficients and effect size for hypothesis testing (based on SmartPLS output)

Hypothesis	Path	ß	T statistics (|O/STDEV|)	P values	f^2	Decision
H4	IQ -> IU	0.001	0.017	0.986	0.000	Rejected
H1	IQ -> UA	0.458	9.768	0.000	0.265	Supported
H2	IQ -> PU	0.264	5.127	0.000	0.104	Supported
H7	PV -> IU	0.038	0.892	0.373	0.003	Rejected
H5	UA -> IU	0.425	8.229	0.000	0.225	Supported
H3	UA -> PU	0.525	11.127	0.000	0.412	Supported
H6	PU -> IU	0.388	6.936	0.000	0.180	Supported

The hypotheses were related to the paths defined in the model (Table 7). The study confirmed significant effects of IQ, UA, and PU on the other factors, but no significant effect was found directly between IQ and IU. In addition, the effect of PV was not proven.

5 Discussion

From a methodological perspective, the model validation and the results confirmed the usability of the simplified model for assessing the intention to use genAI chatbots, namely ChatGPT, in this study. Although model fit indices could be improved, the results are acceptable for a pilot model. The related survey with 13 items may not be exhausting for respondents, and a greater willingness to answer the questions can be expected.

The research question about the impacts of the proposed influencing factors was mostly confirmed. The results show that information quality has a strong influence on the intention to use only through indirect paths. The direct path coefficient is negligible, but perceived usefulness and usability uncover a valuable impact chain. The significant paths explored suggest that the combination of usability and usefulness can explain intention to use remarkably, and the coefficient between them confirms that the perception of usefulness largely depends on the system's usability. The results are in line with former results. Information quality has a strong effect on intermediate constructs (Berényi et al., 2025; Chu 2023; Chen et al., 2024).

The intention to use ChatGPT is high among the respondents, both for study purposes and beyond, and it is not directly dependent on the information quality provided by the system. The intention is based on the perceived usability and usefulness of the users. Information quality and usability explain an acceptable ratio of variance explained (R^2 = 47.2%), and the total model explains 55.9% of the variance in the outcome construct.

According to former studies (Hasanein and Sobaih, 2023; Romero-Rodríguez et al., 2023; Grassini et al., 2024; Gulati et al., 2024; Strzelecki, 2024), behavioral intention significantly predicts genAI use among students, but actual use was not involved in the simplified framework model due to the difficulties of measuring fair values with a self-managed voluntary survey.

The usability construct is a mixed construct within the proposed framework, incorporating facilitating conditions, effort expectancy, and enjoyment. The role is to measure the influencing factors of satisfaction more concisely than the UTAUT2 model.

Regarding effort expectancy, some studies confirmed a positive relationship (Hasanein and Sobaih, 2023), whereas others report non-significant effects, highlighting the context-dependent nature of this construct in genAI adoption. The lack of effect was confirmed by Habibi et al. (2023) and Salifu et al. (2024). The usability construct in this pilot study is a significant contributor.

A notable characteristic of the framework model is that usability incorporates satisfaction. Although it might be worth treating satisfaction as an outcome variable, the opposite approach is intentional since it allows the analysis of its effect in a model where loops and bidirectional relations are not feasible (Hair et al., 2022).

The growing supply of genAI chatbots at various pricing options suggested testing the effect of price-value. The results show that intention to use is not affected by pricing. Price-value was found to be a driver of intention

to use by Romero-Rodríguez et al. (2023), while most studies did not find significant effects in the higher education context. This construct can be omitted if the target audience remains higher education, especially given recent service provider offerings, e.g., Copilot or Gemini access to the academic community.

6 Conclusion

The intention to use genAI chatbots is remarkably high among students at the University of Miskolc, driven primarily by perceived usability and usefulness. At the same time, the direct impact of information quality and price-value is negligible. The results confirmed that promoting cautious use of GenAI is a critical issue. Since genAI may occasionally provide confusing information, rapid, seemingly comprehensive responses can be harmful. The insignificant path between information quality and intention to use is a warning sign, especially since information quality has a stronger effect on usability than on usefulness. That implies that students base their decisions on convenience rather than quality, which poses an educational policy risk. Therefore, university-level policymaking should focus less on restrictions and more on developing the skills to use and raising awareness of computer-generated information. Since the simplified framework model allows quick data collection and is not sensitive to the chatbot used by students, repeated surveys support evidence-based refinement of policies over time.

Despite the contributions and the careful research design, the study has several limitations. Data collection for the pilot study was limited to a single university; further research can test validity using a broader sample. The sample size was acceptable for model testing, but it was not representative. Additionally, existing models were used in the survey design, and simplifications were made on a trial basis. Actual use is not included in the model, which is essential in some situations, given local data and opportunities.

7 Ethics Declaration

The study was conducted in accordance with the Declaration of Helsinki, and approved by the Scientific Committee of the University of Miskolc (TNI/1836/2025, 23. 09. 2025).

8 AI Declaration

No AI tool was used in the development and writing of this paper.

References

Annamalai, N., Bervell, B., Mireku, D.O., and Andoh, R.P.K. (2025) Artificial Intelligence in Higher Education: Modelling Students' Motivation for Continuous Use of ChatGPT based on a Modified Self-determination Theory, *Computers and Education: Artificial Intelligence*, Vol 8, 100346. https://doi.org/10.1016/j.caeai.2024.100346

Ayoubi, K. (2024) Adopting ChatGPT: Pioneering a New Era in Learning Platforms, *International Journal of Data and Network Science*, Vol 8, No. 2, 1341–1348. https://doi.org/10.5267/j.ijdns.2023.11.001

Berényi, Cs. and Csiszárik-Kocsir, Á. (2025) Mesterséges intelligencia a mindennapokban: alkalmazási területek és generációs sajátosságok egy primer kutatás eredményei alapján, *Biztonságtudományi Szemle [Safety and Security Sciences Review]*, Vol 7, No. 3. 179–191. https://doi.org/10.12700/btsz.2025.7.3.179

Berényi, L., Lázár, E., Madaras, Sz. (2025) ChatGPT Acceptance Among Students: An Information System View, *Education Sciences*, Vol 15, No. 11. 1464. https://doi.org/10.3390/educsci15111464

Chen, H.J., Chang, S.T., Chou, P.Y., Tsai, Y.S., Chu, C., Hsieh, F.C. and Tseng, G. (2024) Exploring Users' Continuous Use Intention of ChatGPT Based on the IS Success Model and Technology Readiness, *International Journal of Management Studies and Social Science Research*, Vol 6, No. 1, 366–377. https://doi.org/10.56293/IJMSSSR.2024.4833

Chu, M.N. (2023) Assessing the Benefits of ChatGPT for Business: An Empirical Study on Organizational Performance, *IEEE Access*, Vol 11, 76427–76436. https://doi.org/10.1109/ACCESS.2023.3297447

Davis, F.D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology, *MIS Quarterly*, Vol 13, No. 3, 319. https://doi.org/10.2307/249008

DeLone, W.H. and McLean, E.R. (1992) Information Systems Success: The Quest for the Dependent Variable, *Information Systems Research*, Vol 3, No. 1, 60–95. https://doi.org/10.1287/isre.3.1.60

DeLone, W.H. and McLean, E.R. (2003) The DeLone and McLean Model of Information Systems Success: A Ten-Year Update, *Journal of Management Information Systems*, Vol 19, No. 4, 9–30. https://doi.org/10.1080/07421222.2003.11045748

Farrokhnia, M., Banihashem, S.K., Noroozi, O. and Wals, A. (2024) A SWOT Analysis of ChatGPT: Implications for Educational Practice and Research, *Innovations in Education and Teaching International*, Vol 61, No. 3, 460–474. https://doi.org/10.1080/14703297.2023.2195846

Freeman, J. (2025) „Student Generative AI Survey 2025", [online], HEPI Policy Note 61. Higher Education Policy Institute. https://www.hepi.ac.uk/reports/student-generative-ai-survey-2025/

Grassini, S., Aasen, M.L. and Møgelvang, A. (2024) Understanding University Students' Acceptance of ChatGPT: Insights from the UTAUT2 Model, *Applied Artificial Intelligence*, Vol 38, 2371168. https://doi.org/10.1080/08839514.2024.2371168.

Gulati, A., Saini, H., Singh, H. and Kumar, V. (2024) Enhancing Learning Potential: Investigating Marketing Students' Behavioral Intentions to Adopt ChatGPT, *Marketing Education Review*, Vol 34, No. 3, 201–234. https://doi.org/10.1080/10528008.2023.2300139

Habibi, A., Muhaimin, M., Danibao, B.K., Wibowo, Y.G., Wahyuni. S. and Octavia, A. (2023) ChatGPT in Higher Education Learning: Acceptance and Use, *Computers and Education: Artificial Intelligence*, Vol 5, 100190. https://doi.org/10.1016/j.caeai.2023.100190

Hair, J.F., Hult, G.T.M., Ringle, C.M. and Sarstedt, M. (2022) *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM). Third edition*. SAGE, New York.

Hasanein, A.M. and Sobaih, A.E.E., (2023) Drivers and Consequences of ChatGPT Use in Higher Education: Key Stakeholder Perspectives, European Journal of Investigation in Health, *Psychology and Education*, Vol 13, no. 11, 2599–2614. https://doi.org/10.3390/ejihpe13110181

Hussein, H., Gordon, M., Hodgkinson, C., Foreman, R. and Wagad, S. (2025). ChatGPT's Impact Across Sectors: A Systematic Review of Key Themes and Challenges, *Big Data and Cognitive Computing*, Vol 9, No. 3, 56. https://doi.org/10.3390/bdcc9030056

Jahani, M., Baruah, B. and Ward, A. (2024), „Exploring ChatGPT Utilization Among Master's Students in Higher Education", Paper read at 21st International Conference on Information Technology Based Higher Education and Training (ITHET), IEEE, Paris, France. https://doi.org/10.1109/ithet61869.2024.10837601.

Kock, N. and Hadaya, P. (2018) Minimum sample size estimation in PLS-SEM: The Inverse Square Root and Gamma-exponential Methods, *Information Systems Journal*, Vol 28, No. 1, 227–261. https://doi.org/10.1111/isj.12131

Kovari, A. (2025) Ethical Use of ChatGPT in Education - Best Practices to Combat AI-induced Plagiarism, *Frontiers Education*, Vol 9, 1465703. https://doi.org/10.3389/feduc.2024.1465703

Kumar, A., Kumar, A., Bhoyar, S. and Mishra, A.K. (2024) Does ChatGPT Foster Academic Misconduct in the Future?, *Public Administration and Policy: An Asia-Pacific Journal*, Vol 27, No. 2., 140–153. https://doi.org/10.1108/pap-05-2023-0061

Medina, W.A.A., Velásquez, C.H.F., Zegarra, S.D.R.O., Romero, G.P.M., Andía, A.Q., Bellido, I.E.A., Huaraca, C.R.P., Trejo, F.H.A. and Bravo, N.A. (2024) Using ChatGPT in University Academic Writing: A Bibliometric Review Study on the Implications for Writing Reports, Papers, Essays, and Theses, *International Journal of Learning, Teaching and Educational Research*, Vol 23, No. 12, 360–381. https://doi.org/10.26803/ijlter.23.12.19

Mohamed, A.M., Shaaban, T.S., Bakry, S.H., Guillén-Gámez, F.D. and Strzelecki, A. (2025) Empowering the Faculty of Education Students: Applying AI's Potential for Motivating and Enhancing Learning, *Innovative Higher Education*, Vol 50, 587–609. https://doi.org/10.1007/s10755-024-09747-z

Ringle, C.M., Wende, S. and Becker, J.M. (2024) "SmartPLS 4". [Computer Software], SmartPLS, Bönningstedt.

Romero-Rodríguez, J.M., Ramírez-Montoya, M.S., Buenestado-Fernández, M. and Lara-Lara, F. (2023) Use of ChatGPT at University as a Tool for Complex Thinking: Students' Perceived Usefulness, *Journal of New Approaches in Educational Research*, Vol 12, 323–339. https://doi.org/10.7821/naer.2023.7.1458

Salifu, I., Arthur, F., Arkorful, V., Abam Nortey, S. and Solomon Osei-Yaw, R. (2024) Economics Students' Behavioural Intention and Usage of ChatGPT in Higher Education: a Hybrid Structural Equation Modelling – Artificial Neural Network Approach, *Cogent Social Sciences*, Vol 10, 2300177. https://doi.org/10.1080/23311886.2023.2300177

Strzelecki, A. (2024) To Use or not to Use ChatGPT in Higher Education? A Study of Students' Acceptance and Use of Technology, *Interactive Learning Environments*, Vol 32, No. 9, 5142–5155. https://doi.org/10.1080/10494820.2023.2209881

Venkatesh, V. and Bala, H. (2008). Technology Acceptance Model 3 and a Research Agenda on Interventions, *Decision Sciences*, Vol 39, No. 2, 273–315. https://doi.org/10.1111/j.1540-5915.2008.00192.x

Venkatesh, V., Morris, M.G., Davis, G.B. and Davis, F.D. (2003). User Acceptance of Information Technology: Toward a Unified View, *MIS Quarterly*, Vol 27, No. 3, 425. https://doi.org/10.2307/30036540

Venkatesh, V., Thong, J.Y.L. and Xu, X. (2012) Consumer Acceptance and Use of Information Technology: Extending the Unified Theory of Acceptance and Use of Technology, *MIS Quarterly*, Vol 36, No. 1, 157. https://doi.org/10.2307/41410412

Yang, H. (2024) Towards Responsible Use: Student Perspectives on ChatGPT in Higher Education, *Proceedings of ECEL 23 European Conference on Elearning*, Vol 2023, 415–422. https://doi.org/10.34190/ecel.23.1.2790

Zábrátzky, É. (2025) Az MI oktatásának szükségessége. *Különleges Bánásmód Interdiszciplináris folyóirat [Special Treatment Interdisciplinary Journal]*, Vol 11, No. 2, 65–76. https://doi.org/10.18458/KB.2025.2.65

Zhou, J., Müller, H., Holzinger, A. and Chen, F. (2024) Ethical ChatGPT: Concerns, Challenges, and Commandments, *Electronics*, Vol 13, No. 17, 3417. https://doi.org/10.3390/electronics13173417

Advancing Foresight and Futures Literacy by Generative Artificial Intelligence

Tamás Kristóf
Institute of Entrepreneurship and Innovation, Corvinus University of Budapest, Hungary

Abstract

In the era of accelerating technological change and systemic uncertainty, foresight has become a critical capability for organizations seeking to anticipate and navigate complex futures. At the same time, advances in generative artificial intelligence (GenAI) are rapidly transforming how foresight is conducted, enabling new forms of future exploration, data synthesis, scenario generation, and forward-looking analysis. Despite this growing relevance, the role of GenAI in foresight remains conceptually underdeveloped, with existing research predominantly framing artificial intelligence (AI) as a tool for augmenting human cognition. This paper addresses this gap by conceptualizing the role of GenAI in reshaping foresight practices and contributes to the development of futures literacy – the capacity to understand, imagine, and use the future in decision-making. Drawing on insights from futures studies, knowledge management, and socio-technical systems theory, this paper develops a conceptual framework that positions GenAI as a knowledge infrastructure mediating foresight processes. The paper's main contribution is the identification of a dual-path transformation in AI-supported foresight. The first, the augmentation paradigm, conceptualizes GenAI as a cognitive collaborator that enhances human sensemaking, supports participatory foresight processes, and fosters futures literacy as a learning-oriented capability. The second, the autonomous foresight paradigm, reflects an emerging trajectory in which GenAI systems increasingly perform foresight tasks with limited human intervention, prioritizing efficiency, scalability, and accessibility. The paper highlights a fundamental tension between these two paradigms. While automation expands access to foresight and reduces costs, it may also weaken the reflexive and participatory processes that underpin futures literacy. This raises critical questions regarding the relationship between foresight production and anticipatory learning, as well as the role of human agency in AI-mediated knowledge creation. The paper contributes to the literature by advancing a novel conceptualization of AI-supported foresight as a socio-technical transformation characterized by competing epistemic logics. It further outlines implications for organizational practice, emphasizing the need for hybrid foresight models that integrate AI-driven capabilities with human-centered sensemaking, and identifies key directions for future research on governance, ethics, and human–AI co-creation in foresight.

Keywords: Futures Studies, Foresight, Futures Literacy, Generative AI

Paper type: Academic Research Paper

1 Introduction

In the era characterized by accelerating technological change, systemic uncertainty, and increasingly complex global challenges, organizations face growing pressure to anticipate and navigate uncertain futures. Foresight has therefore emerged as a critical capability for strategic decision-making, enabling organizations to explore alternative futures, identify emerging risks and opportunities, and support long-term resilience. Within this context, futures literacy – the capacity to understand, imagine, and utilize the future – has gained prominence as a key competence for sustainable organizations (Miller, 2018).

Traditionally, foresight practices have relied on expert judgment, stakeholder engagement, qualitative sensemaking, and participatory approaches, such as horizon scanning, trend analysis, technological assessment, Delphi, futures wheel, causal layered analysis, scenario planning, visioning, backcasting, roadmapping, simulations, or gamification. However, recent advances in GenAI are rapidly transforming these practices. Large language models (LLM) and multimodal generative systems are now capable of synthesizing vast amounts of information, generating narrative-rich scenarios, and supporting interactive exploration of uncertainty (Finkenstadt et al., 2024). As a result, GenAI is increasingly positioned as a powerful enabler of foresight, with the potential to enhance anticipatory intelligence and democratize access to futures thinking (Zoccarato et al., 2025).

Despite this growing interest, the role of GenAI in foresight remains conceptually underdeveloped. Existing literature predominantly frames AI as a tool for augmenting human cognition (Mollick, 2024). This aligns with a learning-oriented perspective of foresight, wherein the principal objective is not precision but the augmentation of organizational competence to navigate uncertainty (Poli, 2019). While this perspective captures important developments, it overlooks a more fundamental transformation: the possibility that GenAI may not only support foresight but increasingly perform it autonomously. Emerging applications suggest that generative systems can

execute substantial parts of foresight activities – including horizon scanning, scenario generation, and evaluation – raising important questions about the future role of human agency in anticipatory processes. This shift suggests that GenAI does not merely enhance foresight practices but may fundamentally transform the epistemic logic through which futures are constructed and legitimized.

This paper addresses this gap by proposing a dual-path conceptualization of GenAI-supported foresight, distinguishing between an *augmentation paradigm* and an *autonomous paradigm*. In the augmentation paradigm, GenAI functions as a cognitive collaborator, enhancing human sensemaking and supporting the development of futures literacy. In contrast, the autonomous paradigm reflects an emerging trajectory in which GenAI systems act as semi-autonomous foresight agents, prioritizing efficiency and scalability but potentially altering the epistemic and learning foundations of foresight practice. Emerging market solutions (i.e. https://preempt.life) offering low-cost, AI-generated foresight outputs further indicate a shift toward increasingly automated foresight services.

Building on insights from futures studies, knowledge management, and socio-technical systems theory, this paper develops a conceptual framework that positions GenAI as a knowledge infrastructure mediating these two pathways. It advances a novel theoretical perspective by conceptualizing AI-supported foresight as a dual-path transformation characterized by a tension between augmentation and autonomy. Furthermore, it integrates futures literacy into this framework, highlighting how different modes of human–AI interaction can influence anticipatory learning.

The remainder of the paper is structured as follows. Section 2 reviews the relevant literature on foresight, futures literacy, and GenAI. Section 3 presents the conceptual framework. Section 4 discusses the implications of the dual-path model for theory and practice. Section 5 concludes with key insights and directions for future research.

2 Literature Review

2.1 From prediction to futures literacy

Foresight research has increasingly moved beyond predictive approaches toward a knowledge-based and learning-oriented understanding of the future. Rather than predicting singular outcomes, foresight is conceptualized as a process of sensemaking, anticipation, and knowledge creation under uncertainty. In this context, futures literacy refers to the capability to explore multiple plausible futures, recognize underlying assumptions, and use anticipatory thinking to inform present action (Miller, 2018).

This epistemological shift reflects a broader transition in futures studies toward plurality, reflexivity, and constructivism. Foresight methods and practices are increasingly understood as cognitive and social processes shaping collective organizational intelligence, rather than merely analytical tools. AI-supported foresight further reinforces this perspective by enabling new forms of anticipatory knowledge production (Chodakowska et al., 2026).

2.2 AI as a knowledge infrastructure for foresight

The integration of AI into foresight represents a significant evolution in anticipatory practices. Earlier applications of AI focused primarily on data-driven analysis, including trend identification, clustering, synthesization, and technological assessment. While these approaches enhanced analytical capabilities, they remained largely complementary to expert-driven foresight processes.

The emergence of GenAI introduces a qualitative shift. Unlike traditional analytical tools, GenAI systems – especially LLM – are capable of generating narratives, simulating future scenarios, and engaging in dialogical interactions. This positions GenAI as an active participant in knowledge creation and sensemaking, rather than a passive analytical instrument (Trujillo-Cabezas, 2025; Zoccarato et al., 2025).

Prior empirical research demonstrates that GenAI can significantly enhance foresight practices by enabling rapid scenario generation, synthesis of diverse knowledge domains, and the articulation of complex systemic dynamics (Ferrer i Picó et al., 2025). These developments suggest that GenAI is evolving into a knowledge infrastructure for foresight, fundamentally transforming how futures are explored and constructed.

2.3 Human–AI collaboration in foresight

A central theme in recent literature is the emergence of human–AI collaboration as the prevalent driving force in AI-supported foresight. Rather than replacing human expertise, GenAI is increasingly conceptualized as augmenting cognitive processes and supporting strategic imagination (Hao et al., 2024).

Zoccarato et al. (2025) identify three modes of interaction with GenAI – delegation, confirmation, and cognitive support – each influencing the diversity and critical depth of foresight outcomes. Similarly, Picavet et al. (2025) show that while AI enhances efficiency and scenario richness, effective foresight requires human validation, iterative design, and stakeholder engagement.

Conceptual frameworks further emphasize the importance of collaborative intelligence. Wang (2025) proposes a multi-agent AI architecture that supports foresight processes while maintaining human agency and ethical oversight. In educational contexts, the integration of GenAI with foresight methodologies enhances intellectual agility, creativity, and future-oriented thinking, particularly through iterative and dialogical engagement (Patel & Lim, 2025).

These findings collectively suggest that the value of GenAI lies not in automation but in augmenting human cognition and expanding the space of possible futures.

2.4 Methodological innovation in AI-supported foresight

GenAI is driving substantial methodological innovation in futures studies. Its capacity to generate, iterate, and visualize scenarios enables new hybrid approaches that integrate traditional foresight methods with computational capabilities (Spaniol & Rowland, 2023). For example, hybrid frameworks combining design futures and GenAI enable the creation of archetypal scenarios and participatory tools that enhance engagement and creativity (Bevolo & Draeger, 2025). Similarly, GenAI has been applied to evaluate foresight projects, offering new possibilities for retrospective learning and validation (McIntosh, 2026).

In qualitative research, GenAI functions as an anticipatory infrastructure, supporting the generation of narratives, images, and speculative artifacts that stimulate imagination and facilitate participatory foresight processes (Pink et al., 2025; Gomez-Hernandez, 2026). These developments highlight a shift toward more interactive, experiential, and democratized foresight approaches.

2.5 Critical perspectives: Risks, bias, and reflexivity

Despite its transformative potential, the integration of GenAI into foresight raises important epistemic and ethical concerns. One key issue is the tendency of generative systems to reproduce dominant narratives and biases embedded in their training data, potentially limiting the diversity of imagined futures (Bechthold, 2026). Hence, GenAI might steer futures studies back toward prediction, as opposed to modern human-centered foresight, since it fundamentally analyzes accessible information.

Moreover, the increasing reliance on GenAI in anticipatory governance introduces risks related to opacity, loss of accountability, and technocratic decision-making, particularly when AI-generated outputs are treated as authoritative (Cugurullo & Xu, 2025). These concerns are compounded by the potential erosion of reflexivity if AI-generated futures are accepted uncritically (Floridi, 2023).

Scholars also emphasize the importance of preserving human creativity and embodied sensemaking, warning that over-reliance on AI-generated knowledge may lead to epistemic stagnation and reduced novelty (Peschl, 2024).

2.6 Research gaps

The literature reveals several important gaps. First, there is limited integration between futures theory and GenAI, with most studies focusing on specific applications rather than broader capability development. Second, a multi-level perspective linking individual, organizational, and ecosystem dynamics remains underdeveloped. Third, there is a lack of governance frameworks for responsible AI-supported foresight. Fourth, more research is needed on human–AI co-creation processes and their impact on the quality and diversity of generated futures.

3 Conceptual Framework

The proposed conceptual framework positions GenAI as a knowledge infrastructure that reconfigures the epistemic foundations of foresight, rather than merely a technological tool. Building on insights from futures studies, knowledge management, and socio-technical systems theory, the framework conceptualizes foresight as a multi-layered process of knowledge creation, interpretation, and anticipatory learning.

At the input level, foresight processes draw upon diverse forms of knowledge and contextual information, including weak signals, emerging trends, expert knowledge, stakeholder perspectives, and organizational strategic intent (Chodakowska et al., 2026; Zoccarato et al., 2025). These inputs represent both structured and unstructured knowledge sources that must be synthesized and interpreted under conditions of uncertainty. In

this context, GenAI introduces new capabilities for handling complexity by rapidly integrating heterogeneous data, identifying patterns, and generating coherent representations of possible futures.

At the technological layer, GenAI systems provide a set of core capabilities that fundamentally reshape foresight practices. These include narrative generation, scenario simulation, pattern synthesis, and dialogical interaction (Ferrer i Picó et al., 2025; Trujillo-Cabezas, 2025). Unlike traditional analytical tools, GenAI is capable of producing context-sensitive, narrative-rich outputs and engaging in iterative exchanges with users. As such, it operates not merely as an analytical instrument but as an active participant in sensemaking processes, enabling the co-construction of futures.

A central contribution of the framework is the conceptualization of human–AI interaction as a continuum rather than a fixed mode. In the dominant configuration identified in literature, GenAI functions as a cognitive collaborator within a human-centered foresight process. In this *augmentation paradigm*, interaction modes include delegation of sub-tasks, iterative co-creation, and critical interrogation of AI-generated outputs (Zoccarato et al., 2025; Wang, 2025). Here, foresight remains a fundamentally human-driven activity, with AI augmenting cognitive capacities by expanding the space of possible futures, challenging assumptions, and supporting exploratory thinking.

Within this augmentation pathway, GenAI enhances key foresight processes (Bevolo & Draeger, 2025; Patel & Lim, 2025). By lowering the cost and increasing the speed of scenario generation, GenAI contributes to the democratization of foresight practices, making them more accessible beyond expert communities. Importantly, this pathway supports the development of futures literacy by fostering reflexivity, critical thinking, and the ability to engage with multiple plausible futures (UNESCO, 2023). In this sense, the primary outcome of AI-supported foresight is the enhancement of organizational and individual capacities to navigate uncertainty.

However, the framework introduces a second, more exploratory and disruptive pathway, referred as the *autonomous paradigm*. In this configuration, GenAI systems move beyond augmentation toward performing substantial parts of the foresight process with limited human intervention. Evidence of this trajectory can be found in studies identifying “full delegation” modes of AI use (Zoccarato et al., 2025), multi-agent foresight architectures capable of integrating data and generating scenarios (Wang, 2025), and early applications of AI in anticipatory governance (Cugurullo & Xu, 2025). In the autonomous pathway, GenAI systems may execute end-to-end foresight activities, including horizon scanning, scenario generation, evaluation, and the preparation of comprehensive foresight studies, effectively acting as semi-autonomous foresight agents. This shift increases efficiency and scalability, enabling rapid and cost-effective foresight production. Emerging market developments, including automated foresight platforms, suggest that such configurations are becoming increasingly feasible in practice.

Nevertheless, the autonomous paradigm entails fundamentally different implications for knowledge creation and futures literacy. While it may enhance efficiency and accessibility, it risks reducing the learning-oriented and reflexive dimensions of foresight, which are central to futures literacy. If foresight outputs are generated and consumed without active human engagement, the process may shift from participatory sensemaking to passive reception. Furthermore, autonomous foresight systems may reinforce dominant narratives embedded in existing data and knowledge, limiting the diversity of imagined futures (Bechthold, 2026), and create an illusion of objectivity that obscures underlying assumptions and value judgments (Floridi, 2023).

To account for these tensions, this framework incorporates a set of moderating factors that influence both pathways. These include data biases, ethical governance, transparency, and the preservation of human agency and reflexivity (Floridi, 2023; Cugurullo & Xu, 2025). These moderators shape not only the quality and diversity of foresight outputs, but also the extent to which AI-supported foresight contributes to meaningful futures literacy.

Taken together, the framework conceptualizes GenAI as enabling a dual-path transformation of foresight. On the one hand, it supports human–AI co-creation, enhancing anticipatory intelligence and futures literacy through interactive and reflexive processes. On the other hand, it opens the possibility of increasingly autonomous foresight systems that prioritize efficiency and scalability but may alter the epistemic and learning foundations of foresight practice.

The theoretical contribution of this framework lies in articulating this tension between augmentation and autonomy as a defining characteristic of AI-enabled foresight. By positioning GenAI as an enabling epistemic infrastructure operating across multiple interaction modes, the framework integrates foresight theory, knowledge management, and AI studies into a unified perspective. Moreover, it highlights the importance of human–AI co-creation as a critical mechanism for preserving the transformative potential of foresight in the age of GenAI, while recognizing the emerging trajectory toward automation as a key area for future research.

Figure 1 illustrates how GenAI operates as a mediating knowledge infrastructure, transforming inputs into foresight processes through human–AI interaction, ultimately enhancing futures literacy, while being shaped by ethical and epistemic moderators.

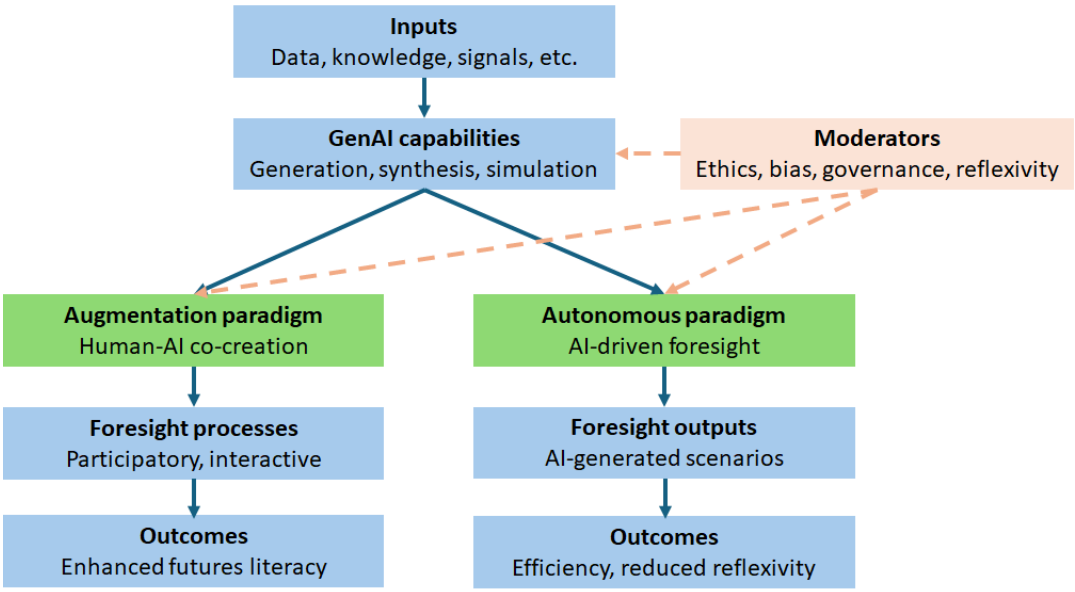


Figure 1. A dual-path conceptual framework of GenAI in foresight
Source: own work

Beyond the distinction between augmentation and autonomy, the proposed framework suggests a deeper shift in the epistemic foundations of futures studies. In the augmentation paradigm, foresight operates as a participatory epistemic process, in which knowledge about the future is co-constructed through interaction, interpretation, and reflexive learning. In contrast, the autonomous paradigm reflects a transition toward a more computational epistemic mode, in which futures are generated through algorithmic synthesis and probabilistic reasoning, with reduced human involvement in meaning-making. This distinction demonstrates that the transformation introduced by GenAI is not merely methodological, but fundamentally epistemological, redefining how futures are known, constructed, and legitimized. Understanding this shift is critical for assessing the long-term implications of AI-supported foresight for organizational learning and strategic decision-making.

Building on this epistemic shift, the framework further implies an anticipatory agency shift, whereby the locus of foresight increasingly moves from human actors toward AI systems. In the augmentation paradigm, agency remains distributed and human-centered, with AI supporting interpretation and exploration. In contrast, the autonomous paradigm reflects a relocation of anticipatory agency toward algorithmic systems, where futures are generated with reduced human intervention. This shift has important implications for accountability, responsibility, and the legitimacy of foresight outputs, particularly in organizational and policy contexts.

4 Discussion

This study sets out to examine how GenAI transforms foresight practices and contributes to the development of futures literacy. Building on the literature and the proposed conceptual framework, the findings suggest that GenAI does not simply enhance existing foresight approaches but introduces a fundamental reconfiguration of the epistemic and organizational foundations of anticipation. This epistemic shift is closely linked to a parallel transformation in anticipatory agency.

4.1. From augmentation to autonomy: A paradigmatic tension

A central insight emerging from this study is the identification of a dual-path transformation in AI-supported foresight.

In the augmentation paradigm, GenAI acts as a cognitive partner that enhances anticipatory intelligence by expanding the space of possible futures, supporting sensemaking, and enabling iterative exploration. This aligns with established views of foresight as a learning-oriented and participatory process (Miller, 2018; UNESCO, 2023), where the primary value lies in developing futures literacy rather than producing accurate predictions.

Here, the interaction between humans and AI fosters reflexivity, critical thinking, and the ability to question underlying assumptions.

In contrast, the autonomous foresight paradigm represents a more disruptive development. As generative systems become capable of performing end-to-end foresight tasks, the role of human actors may shift from active sensemakers to supervisors or consumers of AI-generated outputs. This transition reflects broader trends in AI-enabled knowledge work, where efficiency and scalability increasingly drive adoption.

This duality introduces a fundamental paradigmatic tension. While automation enhances efficiency and accessibility, it may simultaneously undermine the epistemic depth and learning potential that define foresight as a practice. The findings therefore suggest that the future of foresight will not be characterized by a linear progression toward automation, but by the coexistence and interaction of these two paradigms.

4.2. Implications for futures literacy

The distinction between augmentation and autonomy has significant implications for the concept of futures literacy. In the augmentation pathway, futures literacy is actively developed through engagement with uncertainty, exploration of alternative futures, and critical reflection on assumptions. GenAI, in this context, functions as a catalyst for anticipatory learning, enabling individuals and organizations to deepen their understanding of complexity.

However, in the autonomous pathway, the relationship between foresight and learning becomes more ambiguous. If foresight outputs are generated with minimal human involvement, the process may shift from participatory sensemaking to passive consumption. This raises the possibility of a decoupling between foresight production and futures literacy development, where organizations benefit from AI-generated insights without necessarily enhancing their internal anticipatory capabilities.

This distinction highlights a critical question for both research and practice: *Can futures literacy be meaningfully developed without active participation in foresight processes?*

The framework proposed in this paper suggests that while GenAI can increase access to foresight outputs, the transformative potential of futures literacy depends on maintaining human engagement in the process.

4.3. Reconfiguring organizational foresight practices

From an organizational perspective, the integration of GenAI into foresight practices implies a shift from resource-intensive, expert-driven processes toward more distributed, efficient, and technology-mediated approaches. This transformation lowers barriers to entry and enables a wider range of actors to engage with futures thinking.

At the same time, the emergence of autonomous foresight systems raises important questions about the role of expertise, participation, and governance. Traditional foresight practices emphasize stakeholder involvement, collective intelligence, and deliberation. In contrast, AI-driven foresight may centralize knowledge production within technological systems, potentially reducing transparency and inclusiveness.

Organizations therefore face a strategic choice between leveraging GenAI to enhance participatory foresight processes and adopting automated solutions that prioritize efficiency and cost reduction.

Rather than viewing these options as mutually exclusive, the findings suggest that organizations should develop hybrid foresight models, in which AI-driven outputs are integrated into human-centered processes of interpretation and deliberation. Such models can combine the scalability of automation with the reflexivity and contextual sensitivity of human judgment. This tension can be understood as a foresight automation paradox, whereby increasing efficiency through AI may simultaneously reduce the learning-oriented value of foresight.

This suggests that foresight capabilities may increasingly become a strategic differentiator, not only in analytical capacity but in the ability to integrate human and artificial intelligence in anticipatory processes.

4.4. Ethical and epistemic considerations

The increasing role of GenAI in foresight also introduces significant ethical and epistemic challenges. As highlighted in the literature, generative systems are prone to reproducing biases and dominant narratives embedded in their historically available data (Bechthold, 2026), which may constrain the diversity of imagined futures. In autonomous configurations, these biases may become less visible, as the underlying processes of knowledge generation remain opaque.

Moreover, the perceived authority of AI-generated outputs may create an illusion of objectivity, leading decision-makers to over-rely on automated foresight. This risk is particularly pronounced in policy and

governance contexts, where AI systems may influence anticipatory decision-making without adequate transparency or accountability (Cugurullo & Xu, 2025).

These considerations reinforce the importance of embedding ethical governance, transparency, and critical AI literacy within foresight practices. Rather than treating these factors as external constraints, they should be understood as integral components of AI-supported foresight systems, shaping both processes and outcomes.

4.5. Toward a research agenda

The conceptual framework developed in this paper opens several avenues for future research. First, empirical studies are needed to examine how different modes of human–AI interaction influence the quality, diversity, and impact of foresight outcomes. Comparative research between augmentation and autonomous approaches could provide valuable insights into their respective strengths and limitations.

Second, further work is required to explore the relationship between AI-supported foresight and futures literacy. This includes investigating how different levels of human engagement affect learning outcomes at individual and organizational levels.

Third, the development of governance frameworks for responsible AI-supported foresight represents a research priority. Such frameworks should address issues of bias, transparency, accountability, and inclusiveness, ensuring that AI contributes to sustainable and equitable futures.

Fourth, future research should examine the emergence of AI-native foresight practices, including fully automated or multi-agent systems, and their implications for the role of human agency in anticipatory processes.

4.6. Concluding reflection

Overall, this study suggests that GenAI does not simply improve foresight, but it fundamentally redefines its nature. The key challenge is not whether AI can perform foresight tasks, but how its integration reshapes the relationship between knowledge, anticipation, and human agency.

The framework highlights that the future of foresight will likely be shaped by an ongoing tension between augmentation and autonomy. Navigating this tension will be critical for ensuring that the increasing efficiency of AI-supported foresight does not come at the expense of its most valuable outcome: the development of futures literacy as a capability for understanding and shaping complex, uncertain futures.

This tension is particularly relevant in the context of sustainable organizations, where long-term decision-making requires not only efficient foresight outputs but also deep reflexive understanding of complex socio-ecological systems.

5 Conclusions

This paper examined the transformative role of GenAI in foresight practices and its implications for the development of futures literacy. It conceptualized GenAI as a knowledge infrastructure that reshapes how futures are constructed, interpreted, and utilized. In the long term, the evolution of AI-supported foresight may redefine not only how organizations anticipate the future, but who – or what – participates in its construction.

The primary contribution of the paper lies in the development of a conceptual framework, which distinguishes between two emerging paradigms of AI-supported foresight. The augmentation paradigm positions GenAI as a cognitive collaborator that enhances human sensemaking, supports participatory foresight processes, and fosters futures literacy as a learning-oriented capability. The autonomous paradigm reflects an emerging trajectory in which GenAI systems increasingly perform foresight tasks with limited human intervention, prioritizing efficiency and accessibility.

By articulating the tension between augmentation and autonomy, the paper contributes to the literature by reframing AI-supported foresight not as a linear technological progression, but as a socio-technical transformation characterized by competing epistemic logics. While autonomous foresight offers significant advantages in terms of cost reduction and scalability, it also raises concerns regarding reflexivity, diversity of future imaginaries, and the potential decoupling of foresight from futures literacy development. This highlights that the future of foresight will depend not only on technological capabilities, but on how organizations choose to configure the relationship between human and AI.

From a practical perspective, the findings suggest that organizations currently should avoid viewing GenAI as a substitute for human foresight capabilities. Instead, the most promising approach lies in developing hybrid foresight models that combine the efficiency of AI-driven processes with the interpretive and reflective capacities of human actors. Such models can ensure that foresight remains not only a tool for strategic anticipation but also a mechanism for organizational learning and capability building.

However, this study is not without limitations. As a theoretical paper, it cannot provide empirical validation of the proposed framework. Future research is therefore needed to examine how different configurations of human–AI interaction influence foresight outcomes and futures literacy in practice. Comparative studies between augmentation and autonomous approaches, as well as investigations into governance mechanisms for responsible AI-supported foresight, represent promising avenues for further inquiry.

In conclusion, GenAI has the potential to fundamentally redefine foresight. However, the key challenge is not whether AI can perform foresight tasks, but how its integration shapes the relationship between knowledge, anticipation, and human agency. Ensuring that the increasing efficiency of AI-supported foresight does not come at the expense of its most valuable outcome – the development of futures literacy – remains a central challenge for organizations navigating an increasingly complex and uncertain world. Ultimately, the future of foresight will depend not only on what AI can do, but on how organizations choose to engage with it.

6 Ethics declaration

Not applicable.

7 AI declaration

The author employed ChatGPT to enhance the readability and linguistic skill of the paper. The author subsequently reviewed and edited the content and takes full responsibility for this paper's content.

References

Bechthold, L. A. (2026) Rendering new futures or enshrining pictures of the past? The double-edged sword of using AI-generated imagery for anticipation. *Journal of Futures Studies*, In Press. https://jfsdigital.org/rendering-new-futures-or-enshrining-pictures-of-the-past-the-double-edged-sword-of-using-ai-generated-imagery-for-anticipation/

Bevolo, M. and Draeger, D. (2025) Leisure futures: A hybrid approach integrating generative AI and design futures. *Research in Hospitality Management*, Vol. 15, No. 1, pp 77-83. https://doi.org/10.1080/22243534.2025.2450334

Chodakowska, E., Danilczuk, W. and Nazarko, J. (2026) Artificial intelligence assistance in foresight research: Enhancing technology assessment through data-driven methods. *Advances in Science and Technology Research Journal*, Vol. 20, No. 3, pp 299-317. https://doi.org/10.12913/22998624/211285

Cugurullo, F. and Xu, Y. (2025) When AIs become oracles: Generative artificial intelligence, anticipatory urban governance, and the future of cities. *Policy and Society*, Vol. 44, No. 1, pp 98-115. https://doi.org/10.1093/polsoc/puae025

Ferrer i Picó, J., Catta-Preta, M., Trejo Omeñaca, A., Vidal, M., Monguet i Fierro, J. M. (2025) The time machine: Future scenario generation through generative AI tools. *Future Internet*, Vol. 17, 48. https://doi.org/10.3390/fi17010048

Finkenstadt, D. J., Eapen, T. T., Sotiriadis, J. and Guinto, P. (2024) Use GenAI to improve scenario planning. *Harvard Business Review*, Special Issue, Fall, pp 68-71. https://hbr.org/2023/11/use-genai-to-improve-scenario-planning

Floridi, L. (2023) *The ethics of artificial intelligence: Principles, challenges, and opportunities*. Oxford: Oxford University Press.

Gomez-Hernandez, M. (2026) Using generative artificial Intelligence images as a qualitative method to research aging-technology futures. *Qualitative Inquiry*, Online First. https://doi.org/10.1177/1077800425141286

Hao, H., Wang, Y. and Chen, J. (2024) Empowering scenario planning with artificial intelligence: A perspective on building smart and resilient cities. *Engineering*, Vol. 43, No. 12, pp 272-283. https://doi.org/10.1016/j.eng.2024.06.012

McIntosh, A. (2026) Linking generative AI and morphological analysis to conduct foresight evaluation. *Foresight*, Vol. 28, No. 1, pp 66-80. https://doi.org/10.1108/FS-02-2025-0027

Miller, R. (2018) *Transforming the future: Anticipation in the 21st century*. London: Routledge.

Mollick, E. (2024) *Co-intelligence: Living and working with AI*. London: Portfolio/Penguin Random House.

Patel, N. S. and Lim, T-h. J. (2025) Critical design futures thinking and generative AI: A foresight 3.0 approach in higher education to design preferred futures for the industry. *Foresight*, Vol. 27, No. 2, pp 380-402. https://doi.org/10.1108/FS-11-2023-0228

Peschl, M. F. (2024) Human innovation and the creative agency of the world in the age of generative AI. *Possibility Studies & Society*, Vol. 2, No. 1, pp 49-76. https://doi.org/10.1177/27538699241238049

Picavet, M. E. B., Maroni, P., Sandhu, A. and Desouza, K. C. (2025) Human–machine collaboration for strategy foresight: The case of generative AI. *Public Administration Review*, Online First. https://doi.org/10.1111/puar.70048

Pink, S., Korsmeyer, H. and Lyall, B. (2025) Generative AI and broken futures. *Qualitative Inquiry*, Online First. https://doi.org/10.1177/10778004251358070

Poli, R. (2019) *Handbook of anticipation: Theoretical and applied aspects of the use of future in decision making.* Cham: Springer.

Spaniol, M. J. and Rowland, N. J. (2023) AI-assisted scenario generation for strategic planning. *Futures & Foresight Science*, 5, e148. https://doi.org/10.1002/ffo2.148

Trujillo-Cabezas, R. (2025) Exploring the link between foresight and artificial intelligence methods to strengthen collective future-building in contexts of social instability. *Foresight*, Vol. 27, No. 2, pp 267-305. https://doi.org/10.1108/FS-11-2023-0231

UNESCO (2023) *Futures literacy laboratory playbook: An essentials guide for co-designing a lab to explore how and why we anticipate.* Paris: UNESCO & Dhahran: Prince Mohammad Bin Fahd University.

Wang, F. C. (2025) Collaborative foresight in the age of AI: A framework for evolving human-AI dynamics in strategic decision-making and futures research. *Journal of International DBA Studies*, Vol. 1, No. 1, pp 1-28. https://eaj.ebujournals.lu/index.php/JIDS/article/view/155/100

Zoccarato, F., Toletti, G. and Lettieri, E. (2025) Envisioning the future using generative AI – Implications for corporate foresight practices. *Research-Technology Management*, Vol. 68, No. 6, pp 30-41. https://doi.org/10.1080/08956308.2025.2557147

Beyond Solutionism: A Pre-Implementation Diagnostic Framework for Strategic Knowledge Management Alignment

Maayan Nakash
Department of Management, Bar-Ilan University, Ramat Gan, Israel

Abstract

Despite the recognized strategic value of knowledge, many Knowledge Management (KM) initiatives fail due to "solutionism" – the premature deployment of advanced technologies without clarifying business objectives, structural knowledge requirements, or cultural readiness. To address this persistent gap, this paper introduces a pre-implementation diagnostic framework designed to guide Chief Knowledge Officers (CKOs) before designing and deploying Knowledge Management Systems (KMSs). Grounded in the Knowledge-Based View (KBV) and Socio-Technical Systems (STS) theory, the conceptual framework is empirically enriched by participant observation within a large, knowledge-intensive financial institution. The framework is anchored in the central premise that knowledge must be purposefully managed to advance specific organizational goals, thereby positioning a KMS as an enabling socio-technical infrastructure rather than an isolated technological artifact. To operationalize this perspective, the framework prescribes two integrated dimensions: (1) strategic goal setting to align KM initiatives with organizational intent, and (2) a tripartite needs assessment evaluating dynamic knowledge flows, technological specifications, and leadership and cultural commitment. Each dimension is translated into targeted diagnostic questions that convert abstract strategic intent into actionable design requirements and governance practices. By utilizing this diagnostic "compass," organizations can shift from deterministic tool selection to rigorous contextual analysis, identifying structural gaps and mitigating knowledge risks. This ensures that KMS investments generate sustained value and are deeply embedded within the organizational fabric.

Keywords: Knowledge Management Strategy, Pre-Implementation Diagnosis, Strategic Alignment, Knowledge Flow, Knowledge Risk, Socio-Technical Systems

Paper type: Academic Research Paper

1 Introduction

The contemporary volatile and complex business environment compels organizations to leverage knowledge effectively, elevating knowledge management (KM) to a strategic imperative rather than a discretionary function (Bolisani and Bratianu, 2018; Bratianu et al., 2023; do Céu Morais Cláudio and Santos, 2022). KM encompasses systematic processes for creating, structuring, sharing, and applying both explicit assets and tacit know-how to support operational effectiveness and innovation (Alavi and Leidner, 2001). To orchestrate these knowledge flows, organizations deploy knowledge management systems (KMSs) – integrated configurations of digital platforms, governance mechanisms, and work practices. Effective KMSs extend beyond mere storage; they are designed to embed knowledge into everyday workflows and translate dispersed expertise into coordinated organizational action (Thomas, 2024).

Despite sustained scholarly attention and the acknowledgment of knowledge as a critical strategic resource (Bergh et al., 2025), many KM initiatives fail to generate meaningful impact, often resulting in underutilized, symbolic, or abandoned systems (Akhavan et al., 2005; Dalkir, 2013; Haug, 2026; Koivisto and Taipalus, 2025). While prior research has extensively examined KM success factors (Conley and Zheng, 2009; Du Plessis, 2007; Jennex and Olfman, 2009) and developed retrospective maturity models (Escrivão and Silva, 2019; Henz and Oliveira, 2024), comparatively little attention has been devoted to the critical pre-implementation phase. This gap frequently manifests in organizational practice as "solutionism" – the premature deployment of advanced technologies without a clear understanding of the underlying organizational needs. Such a technology-first logic misframes KM as a purely technical challenge (Nakash and Bouhnik, 2022), creating misalignment and yielding repositories that are rich in content but poor in sustained value creation (Davenport and Prusak, 1998).

To address this gap, this study develops a structured pre-implementation diagnostic framework aimed at guiding KM decision-making. The conceptual development of this framework is empirically enriched by prolonged participant observation and ethnographic engagement within a large, knowledge-intensive organization. Designed to counter solutionism and mitigate the risk of implementation failure, the proposed framework equips Chief Knowledge Officers (CKOs) with a diagnostic compass prior to resource allocation. It is structured around two complementary dimensions: 1) strategic goal setting, which anchors KM design in business strategy, and 2) needs assessment, which translates these goals into operational requirements across

knowledge flows, technological readiness, and leadership commitment. Ultimately, this pre-implementation diagnosis ensures that KM investments result in fit-for-purpose systems that are deeply embedded in organizational workflows and sustainable over time.

2 Research Foundations: Theoretical and Field-Based Insights

2.1 Theoretical Background

The proposed diagnostic framework is grounded in complementary streams of organizational research that elucidate why KM initiatives frequently underperform when introduced without systematic prior diagnosis.

2.1.1 The Knowledge-Based View and Strategic Alignment

The Knowledge-Based View (KBV) positions knowledge as the firm's most critical strategic resource, asserting that sustained competitive advantage stems from the organization's capability to integrate, recombine, and apply knowledge in pursuit of its goals (Curado and Bontis, 2006; Eisenhardt and Santos, 2002; Grant, 1996). Within this paradigm, knowledge acquires economic value exclusively when it is embedded within organizational routines, decision-making processes, and coordinated actions. Consequently, knowledge assets that remain disconnected from core activities become inert, regardless of their scope or quality (Davenport and Prusak, 1998; Nakash and Bolisani, 2025; Nonaka and Takeuchi, 1995).

Strategic Alignment Theory complements this by establishing that organizational investments yield performance returns only when coherently linked to strategic intent (Henderson and Venkatraman, 1999). In the KM domain, misalignment occurs when systems are implemented without a clear articulation of how they resolve priority challenges. Under such conditions, KM initiatives may function efficiently at a technical level yet remain strategically irrelevant (do Céu Morais Cláudio and Santos, 2022). This theoretical linkage underscores that the efficacy of a KMS is fundamentally contingent upon its alignment with broader organizational strategies rather than its technological completeness.

2.1.2 The Socio-Technical Nature of Knowledge Management

KM is inherently socio-technical, emerging from the continuous interaction between digital infrastructures and social systems (Assegaff and Hussin, 2012; Dalkir, 2013; Šajeva, 2010). Socio-Technical Systems Theory (STS) emphasizes that organizational effectiveness relies on the joint optimization of technical components (e.g., platforms, tools, formal protocols) and social components (e.g., professional roles, organizational norms, leadership practices, motivational structures) (Cherns, 1976, 1987). From this perspective, a KMS is not a neutral repository; it is an organizational arrangement that both shapes and is shaped by everyday work practices (Nakash and Bouhnik, 2022).

The STS lens explains why technology-centric KM initiatives frequently fail. Systems designed without regard for existing workflows, professional identities, or internal power dynamics often encounter resistance, superficial adoption, or the creation of parallel "shadow" workarounds. These outcomes are manifestations of socio-technical misalignment. Effective KM therefore requires congruence between system affordances and organizational realities, positioning social conditions – such as trust and psychological safety – as constitutive elements of the KMS rather than secondary considerations (Figueiredo et al., 2025; Liebowitz and Beckman, 2020; Min et al., 2022; Thomas, 2024).

2.1.3 Knowledge Flow Dynamics

A further foundation lies in the processual view of knowledge as a dynamic and situated organizational asset (Bratianu, 2023; Bratianu and Bejinaru, 2026). Knowledge generates value through movement – across individuals, departments, and tasks – via interconnected processes of creation, codification, sharing, and application (Durst and Zieba, 2019; Liebowitz and Beckman, 2020; Nonaka and Takeuchi, 1995). Therefore, KM effectiveness is measured by the continuity and reliability of these knowledge flows, not merely by the capacity of storage repositories (Davenport and Prusak, 1998; Jennex and Olfman, 2009).

These flows are frequently disrupted by structural silos, fragmented digital ecosystems, and misaligned incentives (Nakash, 2025). Such challenges are compounded by the fundamental distinction between tacit and explicit knowledge. Tacit knowledge, which is deeply embedded in human experience and judgment, resists rigid codification and relies heavily on interpersonal interaction and trust. Conversely, explicit knowledge benefits from systematic documentation and standardization (Collins, 2019; Nonaka, 1994). Recognizing these fundamental differences is essential to avoid structural mismatches that frequently undermine KM implementations despite heavy technological investments.

2.2 Participant Observation and Contextual Methodology

Complementing its theoretical foundations, the development of the diagnostic framework was informed by prolonged participant observation within a complex, knowledge-intensive organizational setting. This ethnographic engagement spanned 18 months at a major financial institution with approximately 8,000 employees, providing a rich context for observation across diverse divisions, including retail banking, operations, risk management, compliance, and strategy. To ensure methodological rigor and uphold established ethical standards within the scientific community (Lipson, 1994), the research was conducted while strictly maintaining participant anonymity and data confidentiality.

The researcher adopted an embedded observer role, facilitating access to both formal KM initiatives and informal daily work routines through a multifaceted qualitative data collection strategy. This involved sustained field presence and reflective journaling, complemented by ongoing informal and semi-structured conversations with a broad cross-section of the workforce – ranging from CKOs and KM tool implementers to frontline knowledge leaders.

The primary focus of this observation was not to evaluate specific software tools, but rather to examine the pre-implementation ecosystem: specifically, how KM initiatives were conceptualized, justified, translated into system requirements, and ultimately experienced within the organization. The observational data were analyzed using an inductive thematic approach (Clarke and Braun, 2017) to uncover recurring systemic tensions. Ultimately, these field-observed misalignments directly informed the identification of the "diagnostic blind spots" that constitute the structural basis of the proposed pre-implementation framework.

3 Conceptual Anchor

Building upon the theoretical and empirical foundations discussed previously, the proposed diagnostic approach is anchored in a singular conceptual premise: KM must function as a strategic conduit rather than an isolated administrative end. When conceptualized merely as a functional imperative or confined within a technological silo, KM initiatives risk becoming symbolic exercises detached from actual organizational value creation.

Consequently, effective KM necessitates deliberate calibration with specific organizational objectives. Core knowledge processes – spanning acquisition, documentation, sharing, and application – must be explicitly designed to drive tangible outcomes, such as accelerating onboarding efficiency, minimizing operational errors, or enhancing competitive differentiation. Within this paradigm, knowledge derives its value strictly from its capacity to inform action and elevate performance, rather than from its mere codification or storage.

Accordingly, KMSs are conceptualized not as self-contained technological artifacts, but as socio-technical infrastructures that enable knowledge work. Their architecture, scope, and governance must dynamically mirror both strategic intent and operational realities. This pragmatic orientation serves as the interpretive lens for the proposed framework, dictating the translation of abstract strategic aspirations into concrete, actionable system requirements through targeted diagnostic inquiry.

4 The Diagnostic Framework

To operationalize the conceptual anchor, the proposed diagnostic framework is structured around two sequential dimensions:

- The “Why” – Strategic Goal Setting (Section 4.1)
- The “What” and “How” – Needs Assessment (Section 4.2)

This deliberate sequencing ensures that KM initiatives are firmly rooted in strategic intent prior to the translation of broad objectives into specific operational and technical requirements.

4.1 Strategic Goal Setting

Purpose: The initial diagnostic phase is designed to mitigate the risk of "solutionism" – the premature deployment of technology in search of a problem. Its primary objective is to guarantee that KM interventions address critical organizational imperatives rather than degrading into symbolic, isolated IT projects.

Table 1: Guiding Questions for Strategic Goal Setting

Dimension	Guiding Diagnostic Questions
Strategic Necessity	• Is the KM initiative explicitly linked to the organization's overarching strategy? • Why is a KMS required at this specific point in time?
Process Alignment	• Which core business processes will the KMS support? • How will these processes be transformed or enhanced?
Success Metrics	• What Key Performance Indicators (KPIs) will define success? • How will KM's contribution to organizational performance be measured?
Operational Feasibility	• Have employees' daily workflows been analyzed to inform design? • Which goals are prioritized, and what trade-offs (scope, speed, investment) are acceptable?

Rationale: These guiding questions function as a strategic threshold, mandating clarity before any technological resource allocation occurs. KM initiatives frequently originate from vague aspirations, such as "enhancing collaboration", which lack actionable specificity and baseline metrics. If organizational leadership cannot articulate a compelling urgency for the KM investment (e.g., mitigating tacit knowledge loss due to expert retirement, or accelerating time-to-market in R&D), the initiative is highly vulnerable to becoming a superficial undertaking.

Furthermore, this phase forces an explicit alignment with macro-organizational goals; for instance, an enterprise strategy focused on global expansion requires a KMS architecture deliberately optimized for cross-border knowledge sharing. By interrogating how core processes will demonstrably change, the framework compels KM leaders to precisely define the organizational "to-be" state. Ultimately, shifting the discourse from technological implementation to measurable value creation ensures the system is evaluated on its capacity to enable action rather than its mere operational existence.

4.2 Needs Assessment

Once the strategic imperative (the "Why") is firmly established, the diagnostic framework advances to the operational level. The needs assessment phase translates overarching strategic goals into actionable system specifications across three interdependent domains:

- Knowledge Flow Processes (Section 4.2.1)
- Technological Implementation Specification (Section 4.2.2)
- Leadership and Cultural Commitment (Section 4.2.3)

Together, these dimensions provide a holistic evaluation of what specific knowledge is required, how it should circulate throughout the organization, which technical architectures can best enable this movement, and whether the enterprise possesses the cultural and leadership capacity to sustain it.

4.2.1 Knowledge Flow Processes

Purpose: Mapping knowledge flows is essential for understanding the lifecycle of knowledge within the organization – specifically how it is acquired, documented, shared, and applied. Effective KMS design is fundamentally contingent upon identifying structural gaps, characterizing knowledge consumers, pinpointing critical points of application, and diagnosing systemic barriers to knowledge dissemination.

Table 2: Guiding Questions for Knowledge Flow Processes

Dimension	Guiding Diagnostic Questions
General Ecosystem	• Who consumes the knowledge? • Are consumers centralized or dispersed? • What are the main work interfaces?
Knowledge Acquisition	• What are the primary sources? • How reliable are they? • Who originates the knowledge? • What proportion is tacit vs. explicit?
Knowledge Documentation	• Where are the main documentation nodes? • How much knowledge must be captured and in what format? • Who validates and approves content? • Is there a plan for content currency?

Dimension	Guiding Diagnostic Questions
Knowledge Sharing	• How accessible and retrievable is the knowledge? • Is sharing institutionalized or ad hoc? • What are the current channels? • What barriers exist (technical, procedural, cultural)? • Can professional communities be leveraged?
Knowledge Application	• When and how is knowledge needed (real-time vs. on-demand)? • Preferred consumption mode: Push (automated alerts) or Pull (active search)? • Is the knowledge temporary or permanent?

Rationale: Addressing these specific diagnostic questions ensures that KM initiatives accurately reflect the socio-technical realities of the workforce rather than being dictated solely by software capabilities. For example, precise user segmentation is critical; an R&D department and a sales division inherently require different levels of knowledge granularity, rendering a “one-size-fits-all” system architecture highly detrimental. Furthermore, if the diagnosis reveals that the organization's most critical asset is tacit knowledge, a static document repository will inevitably fail. In such scenarios, the architectural solution must purposefully emphasize "connecting people to experts" rather than merely "collecting documents."

Additionally, rigorously assessing knowledge quality and ownership protocols prevents the KMS from degrading into an obsolete repository – a significant knowledge risk that erodes user trust. This analysis also exposes whether current knowledge-sharing behaviors rely on ad hoc individual efforts or systemic planning, thereby enabling targeted managerial interventions to dismantle structural silos and alleviate time constraints. Retention strategies must also be explicitly tailored based on whether the knowledge is tribal (deeply embedded in the minds of specific experts) or ephemeral. Finally, aligning consumption modes with actual workflows is vital for sustained adoption: a frontline support agent, for instance, requires a robust "Pull" mechanism (rapid search during live client interactions), whereas a compliance officer necessitates a reliable "Push" mechanism (automated alerts regarding regulatory shifts).

4.2.2 Technological Implementation Specification

Purpose: Technological readiness is a critical determinant of KM success. This dimension evaluates whether the organization’s existing digital infrastructure can robustly support the defined knowledge goals or if it requires significant augmentation. The overarching objective is to ensure seamless integration into daily workflows, thereby avoiding the creation of fragmented technological silos that undermine system adoption.

Table 3: Guiding Questions for Technological Implementation Specification

Dimension	Guiding Diagnostic Questions
Existing Infrastructure	• Which systems are currently in place (CRM, ERP, SharePoint, Slack)? • Do they enable or conflict with KM processes?
Integration & Interoperability	• Do systems integrate smoothly or operate as isolated silos? • Can knowledge flow automatically between the proposed KMS and core business applications?
Search & Retrieval	• Is there a robust, multi-channel search engine indexing federated repositories? • How effective is retrieval in terms of speed and relevance?
Functional Gaps	• Which knowledge processes are unsupported by current technology?
User Accessibility (UX)	• How easily can employees access systems (Single Sign-On, mobile, remote)? • Is the interface intuitive enough to minimize friction?
Security & Governance	• What security constraints and compartmentalization policies exist? • How do these balance protection with accessibility?

Rationale: Addressing these questions directly mitigates the risk of deploying technological tools that are fundamentally incompatible with the existing digital ecosystem. If a KMS introduces procedural friction, employees will inevitably bypass it. Rigorous integration checks avert the creation of isolated "knowledge islands" where critical knowledge is trapped and rendered unsearchable. Search and retrieval capabilities are equally paramount; without highly intuitive, federated search experiences, user trust in the system erodes rapidly.

Furthermore, security governance must strike an optimal equilibrium – avoiding both excessive compartmentalization, which stifles cross-functional knowledge flow, and unchecked accessibility, which exposes the organization to severe knowledge risks (such as unauthorized exposure or misuse). Ultimately, this

diagnostic specification converts abstract operational needs into concrete technical requirements, ensuring the technology functions as a seamless enabler rather than an administrative barrier.

4.2.3 Leadership and Cultural Commitment

Purpose: This dimension assesses the organization's socio-cultural infrastructure – namely, leadership and organizational culture – which profoundly influences KM outcomes. It is designed to distinguish genuine strategic intent from merely symbolic gestures, evaluating whether existing governance structures, motivational frameworks, and organizational mindsets can sustain a robust knowledge culture.

Table 4: Guiding Questions for Leadership and Cultural Commitment

Dimension	Guiding Diagnostic Questions
Leadership Support	• Is there visible backing from C-level executives? • Are leaders modeling desired behaviors?
Cultural Readiness	• Does the organization foster psychological safety for learning from mistakes? • Is knowledge hoarding rewarded or discouraged?
Incentives & Motivation	• What mechanisms reward knowledge sharing and KMS adoption? • Are incentives intrinsic (recognition) or extrinsic (bonuses), and do they align with team goals?
Sustainability & Vision	• Is KM viewed as a continuous process or a one-time project? • Are resources allocated for long-term maintenance?

Rationale: While technology provides the structural capacity for knowledge exchange, organizational culture dictates the behavioural willingness to engage in it. Visible and sustained leadership sponsorship is pivotal; absent this support, KM is quickly relegated to a discretionary function, rendering it highly vulnerable during budget constraints.

Furthermore, incentive structures must systematically address the individual motivational calculus: if performance metrics disproportionately reward isolated individual achievement, employees lack a rational imperative to share knowledge. By exposing whether the initiative is merely symbolic, this framework forces a critical reality check prior to implementation: are genuine behavioural shifts feasible, or is the organization simply deploying software? These insights are crucial for developing tailored change management strategies to dismantle entrenched cultural barriers.

5 Discussion

This study bridges a critical gap in contemporary KM literature and practice by shifting the analytical focus from post-implementation maturity to pre-implementation diagnosis. While traditional KM maturity models provide valuable descriptive benchmarks for evaluating existing organizational capabilities, they are inherently retrospective (Escrivão and Silva, 2019; Henz and Oliveira, 2024). By proposing a structured diagnostic compass, this research moves beyond the pervasive trap of "solutionism," offering a prescriptive, methodological bridge that connects abstract organizational theories with the localized, operational realities of system deployment prior to implementation.

A primary theoretical contribution of this framework is the reconceptualization of organizational "readiness." Moving away from technologically deterministic models that equate readiness solely with IT infrastructure (Alavi and Leidner, 2001; Thomas, 2024), this approach strictly applies STS theory to the pre-implementation phase. Specifically, the proposed diagnostic dimensions systematically address the four interacting variables central to STS configurations: technology, people, tasks, and structure (Bostrom and Heinen, 1977). The technological specification targets the technology; the assessment of cultural commitment evaluates the people; mapping knowledge flows addresses the tasks; and strategic goal setting aligns with the organizational structure. This mapping highlights that technological maturity alone is entirely insufficient for success; rather, the technological architecture must be jointly optimized with these strategic, cultural, and processual dimensions. By exposing how a lack of psychological safety or a culture that implicitly rewards knowledge hoarding can derail even the most advanced platforms, the framework demonstrates that socio-technical congruence is a fundamental prerequisite for adoption, not merely a secondary enabler.

Furthermore, the framework effectively operationalizes the KBV (Grant, 1996) by translating its abstract principles into an actionable diagnostic protocol. Instead of treating knowledge as a static commodity to be warehoused, the needs assessment focuses on mapping dynamic knowledge flows and rigorously differentiating between tacit and explicit requirements. By interrogating actual work-as-done – rather than idealized, formal

procedures – the framework uncovers critical socio-technical friction points, such as tacit knowledge bottlenecks or structural mismatches in knowledge consumption modes (e.g., Push vs. Pull). This empirical grounding ensures that knowledge assets are deliberately mobilized to enhance performance and inform action, thereby fulfilling the core tenet of the KBV.

From a managerial and implementation standpoint, the diagnostic process functions as a vital mechanism for early intervention and risk mitigation. It equips CKOs with a structured mandate to identify structural silos, evaluate the authenticity of leadership commitment, and calibrate technical specifications before irrecoverable resources are expended. Consequently, this diagnostic phase actively shifts the perception of KMSs from peripheral IT projects to core strategic infrastructures intricately linked to organizational performance outcomes.

Finally, acknowledging the volatility of the contemporary business environment (Bolisani and Bratianu, 2018; Botha et al., 2014; Bratianu et al., 2023), the framework is expressly designed for iterative application. As strategic priorities pivot and digital ecosystems inevitably evolve, the diagnostic tool acts as a continuous compass, providing a current operational bearing while supporting necessary recalibrations. This longitudinal adaptability ensures that KM initiatives remain deeply embedded within the organizational fabric—functioning as resilient, evolving infrastructures rather than finite, standalone projects.

6 Conclusion

Successful KMS implementation fundamentally transcends mere technological readiness; it necessitates deliberate alignment with strategic objectives, optimized knowledge flows, and unwavering cultural commitment. The proposed pre-implementation diagnostic framework provides a structured methodology to achieve this essential socio-technical congruence. By integrating strategic goal setting with a comprehensive, tripartite needs assessment, the framework shifts the organizational focus from deterministic tool selection to context-driven analysis. Rather than prescribing universal solutions, it compels organizations to engage in critical diagnostic inquiry, thereby mitigating the risk of "solutionism" and maximizing long-term value realization. Future research should seek to empirically validate this framework across diverse sectors, specifically exploring the longitudinal and structural relationships between these pre-implementation diagnostic indicators and measurable organizational performance outcomes.

7 Ethics Declaration

The author declares that this study adhered to scientific ethical standards, ensuring strict participant anonymity and data confidentiality.

8 AI Declaration

The author confirms that no AI tools were used in the creation of this manuscript.

References

Akhavan, P., Jafari, M. and Fathian, M. (2005), "Exploring the failure factors of implementing knowledge management system in the organizations", *Journal of Knowledge Management Practice*, Vol.6.

Alavi, M. and Leidner, D.E. (2001), "Knowledge management and knowledge management systems: Conceptual foundations and research issues1, 2", *MIS quarterly*, Vol.25 No.1, pp.107-136.

Assegaff, S. and Hussin, A.R.C. (2012), "Review of knowledge management systems as socio-technical system", *arXiv preprint arXiv:1212.0387*.

Bergh, D.D., D'Oria, L., Crook, T.R. and Roccapriore, A. (2025), "Is knowledge really the most important strategic resource? A meta-analytic review", *Strategic Management Journal*, Vol.46 No.1, pp.3-18.

Bolisani, E. and Bratianu, C. (2018), "Emergent knowledge strategies", *Emergent Knowledge Strategies*, Vol.4, pp.23-47.

Bostrom, R.P. and Heinen, J.S. (1977), "MIS problems and failures: A socio-technical perspective: Part II: The application of socio-technical theory", *MIS quarterly*, Vol.1 No.4, pp.11-28.

Botha, A., Kourie, D. and Snyman, R. (2014), *"Coping with continuous change in the business environment: Knowledge management and knowledge management technology"*, Elsevier.

Bratianu, C. (2023), "Knowledge dynamics: Exploring its meanings and interpretations. Management Dynamics in the Knowledge Economy", Vol.11 No.2, pp.100-111.

Bratianu, C. and Bejinaru, R. (2026), "Knowledge dynamics", In *Encyclopedia of Diversity, Equity, Inclusion and Spirituality* (pp.867-879). Cham: Springer Nature Switzerland.

Bratianu, C., Handzic, M. and Bolisani, E. (Eds.). (2023), "*The future of knowledge management: reflections from the 10th anniversary of the International Association of Knowledge Management (IAKM)*", Vol.12, Springer Nature.

Cherns, A. (1976), "The principles of sociotechnical design", *Human relations*, Vol.29 No.8, pp.783-792.
Cherns, A. (1987), "Principles of sociotechnical design revisted", *Human relations*, Vol.40 No.3, pp.153-161.
Clarke, V. and Braun, V. (2017), "Thematic analysis", *The journal of positive psychology*, Vol.12 No.3, pp.297-298.
Collins, H. (2019), *"Tacit and explicit knowledge*", University of Chicago press.
Conley, C.A. and Zheng, W. (2009), "Factors critical to knowledge management success", *Advances in Developing Human Resources*, Vol.11 No.3, pp.334-348.
Curado, C. and Bontis, N. (2006), "The knowledge-based view of the firm and its theoretical precursor", *International Journal of Learning and Intellectual Capital*, Vol.3 No.4, pp.367-381.
Dalkir, K. (2013), *"Knowledge management in theory and practice*", Routledge.
Davenport, T.H. and Prusak, L. (1998), *"Working knowledge: How organizations manage what they know"*, Harvard Business Press.
do Céu Morais Cláudio, M. and Santos, A. (2022), "Strategic Alignment of Knowledge Management Systems", In *International Conference on Technology and Innovation in Learning, Teaching and Education* (pp.418-433), Cham: Springer Nature Switzerland.
Du Plessis, M. (2007), "Knowledge management: what makes complex implementations successful?", *Journal of Knowledge Management*, Vol.11 No.2, pp.91-101.
Durst, S. and Zieba, M. (2019), "Mapping knowledge risks: towards a better understanding of knowledge management", *Knowledge Management Research & Practice*, Vol.17 No.1, pp.1-13.
Eisenhardt, K.M. and Santos, F.M. (2002), "Knowledge-based view: A new theory of strategy", *Handbook of strategy and management*, Vol.1 No.1, pp.139-164.
Escrivão, G. and Silva, S.L.D. (2019), "Knowledge management maturity models: Identification of gaps and improvement proposal", *Gestão & Produção*, Vol.26, e3890.
Figueiredo, P., Rodrigues, R. and Diogo, A. (2025), "From safety to sharing: A bibliometric mapping of psychological safety, knowledge management, and organizational learning", *Administrative Sciences*, Vol.15 No.11, 427.
Grant, R.M. (1996), "Toward a knowledge-based theory of the firm", *Strategic management journal*, Vol.17 No.S2, pp.109-122.
Haug, A. (2026) "Critical success factors for knowledge management systems in new product development", *Knowledge Management Research & Practice*, pp.1-13.
Henderson, J.C. and Venkatraman, H. (1999), "Strategic alignment: Leveraging information technology for transforming organizations", *IBM systems journal*, Vol.38 No.2.3, pp.472-484.
Henz, J.L. and Oliveira, M. (2024), "Knowledge management implementation: A systematic literature review", *Knowledge and Process Management*, Vol.31 No.4, pp.284-294.
Jennex, M.E. and Olfman, L. (2009), "A model of knowledge management success", In *Selected Readings on Information Technology Management: Contemporary Issues* (pp.76-93), IGI Global Scientific Publishing.
Koivisto, K. and Taipalus, T. (2025), "Pitfalls in effective knowledge management: insights from an international information technology organization", *International Journal of Knowledge Management Studies*, Vol.16 No.1, pp.27-60.
Liebowitz, J. and Beckman, T. J. (2020), *"Knowledge organizations: What every manager should know"*, CRC Press.
Lipson, J.G. (1994), "Ethical issues in ethnography", *Critical issues in qualitative research methods*, pp.333-355.
Min, J., Lee, J., Ryu, S. and Lee, H. (2022), "The effects of interaction between team climates and KMS value perception on knowledge activities: a multilevel socio-technical systems approach", *Information Technology and Management*, Vol.23 No.1, pp.1-21.
Nakash, M. (2025), "Sharing as Power: Keys Strategies for Breaking Knowledge Silos", In *European Conference on Knowledge Management*, Vol.1, pp.703-708, Academic Conferences International Limited.
Nakash, M. and Bolisani, E. (2025), "Making knowledge management transparent: a new perspective on KM processes integration in the organizational framework", *Business Process Management Journal*, Vol.31 No.8, pp.49-66.
Nakash, M. and Bouhnik, D. (2022), "'A system that will do magic': organizational perspective on the technological layer in knowledge management", *Aslib Journal of Information Management*, Vol.74 No.6, pp.1089-1102.
Nonaka, I. (1994), "A dynamic theory of organizational knowledge creation", Organization science, Vol.5 No.1, pp.14-37.
Nonaka, I. and Takeuchi, H. (1995), *"The knowledge-creating company"*, Oxford University Press.
Šajeva, S. (2010), "The analysis of key elements of socio-technical knowledge management system", *Economics & Management*.
Thomas, A. (2024), "Digitally transforming the organization through knowledge management: A socio-technical system (STS) perspective", *European Journal of Innovation Management*, Vol.27 No.9, pp.437-460.

Emotional Knowledge and Manipulative Design in AI-Powered Fashion Recommendation Systems

Hairong Zhai, Farag Edghiem
Faculty of Arts and Humanities, Manchester Metropolitan University, Manchester, United Kingdom

Abstract

AI-powered recommendation systems have evolved to offer more than just convenience, increasingly shaping consumer identity and purchasing behaviour in the online fashion retail sector. In the context of fashion e-commerce, recommendation systems assist users in navigating information overload by predicting their preferences and suggesting items that align with their tastes. Technical advances, such as deep learning and multimodal integration, significantly improved the performance of fashion recommendation systems. These systems can now incorporate product images, textual descriptions, reviews, as well as videos to capture style, fit, and context more precisely. Nevertheless, these technical advances that could enhance accuracy also empower greater influence over how consumers perceive themselves, others, and what is considered desirable.

Extant research has predominantly focused on technical search engine optimisation and algorithmic performance, leaving the ethical, psychological, and emotional dimensions of fashion recommendation comparatively neglected, particularly significant given fashion's intimate links to self-image, body perception, and social validation. Online interfaces routinely employ dark patterns such as scarcity cues, obstruction, and interface interference to exploit cognitive biases and bounded rationality. Users frequently feel manipulated even when aware of it, expressing frustration and resignation toward persuasive design. Manipulation and autonomy have been identified as central ethical challenges in AI-based systems, alongside concerns about the algorithmic exploitation of vulnerabilities such as insecurity and compulsive tendencies. These debates have, however, largely overlooked the fashion-specific context and the tacit, emotional, and relational knowledge that such systems produce, reshape, and exploit.

This paper contributes towards addressing the paucity of research on emotional fashion recommendation by examining the implementation of manipulative design strategies, that is integrated into AI-powered fashion recommendation systems, and how users experience them. This study conceptualises fashion recommendation systems as socio-technical, arguing that they manage not only data, but also tacit, emotional and relational knowledge. Six semi-structured interviews were conducted with Generation Z fashion consumers in the United Kingdom with data analysed thematically to explore emotional and subjective responses to algorithmic persuasion, urgency cues, social validation signals, and curated recommendation feeds. The study findings make the following contributions: it extends dark pattern and algorithmic persuasion research into the under-explored domain of fashion; it foregrounds the emotional dimensions of manipulation, showing how aesthetic and social cues shape identity-related behaviour beyond cognitive bias; and it offers practical ethical insights for designers, retailers, and policymakers developing more transparent, responsible, and human-centred recommendation systems.

Keywords: AI-powered Recommendation Systems, AI-driven Fashion Retail, Manipulative Design and Dark Patterns, Digital Ethics, Tacit and Emotional Knowledge

Paper type: Academic Research Paper

1 Introduction

It has been frequently debated that when consumers mention a product in conversations, their social media feeds ‘ For You Page’ flood with related advertising content. More persistently when searching for fashion items algorithmic suggestions carry on across platforms. These increasingly recognised digital experiences are powered by AI-driven recommendation systems, technologies designed to track user behaviour and deliver personalised suggestions. Recommendation systems use complex algorithms to process extensive amounts of user data, helping individuals navigate overwhelming product landscapes and discover items aligned with their tastes and interests (Masciari et al., 2024).

In the fashion domain, recommendation systems play a particularly persuasive role. They not only suggest products based on an individual’s browsing and purchasing history, but also factor in social signals, aesthetic preferences, and real-time trend analysis. Social-network-based fashion recommendation models consider factors such as a user’s social circle, fashion product attributes, and stylistic suggestions that make consumer feel personally and socially validated (Chakraborty et al., 2021). The economic benefits for fashion retailers are clear, as these algorithms enhance customer engagement, increase sales, and provide a competitive edge.

However, over-reliance on recommendation systems raises growing concerns. Previous studies show that nearly all users are aware of personalised recommendations and acknowledge difficulty resisting their influence,

even when they are aware of manipulative design patterns (Bongard-Blanchy et al., 2021; Mathur et al., 2019). While there is limited academic work clearly linking manipulation and fashion recommendation systems, this study aims to critically examine the manipulative design strategies employed by AI-powered fashion recommendation systems to influence consumer behaviour, and to explore the ethical implications of fashion recommendation systems for consumer autonomy.

2 Theoretical Background

2.1. The Evolution and Application of AI-Based Recommendation Systems

Recommendation systems have become an essential component of modern e-commerce, offering a solution to the "information overload" problem faced by consumers (Geetha & Renuka, 2019). A recommendation system is to predict a user's preference for a set of items, such as movies, music, or fashion items. Early recommendation methods relied on collaborative filtering and content-based filtering, but recent advancements in AI, particularly the application of deep learning models, have significantly enhanced their accuracy and scalability. Geetha and Renuka (2019) note that deep learning, a sub-domain of machine learning, is utilised in RSs to enhance the feature of predicted outputs. Similarly, Chakraborty et al. (2021) highlight the use of deep matrix factorisation models and neural autoregressive approaches in collaborative filtering for fashion, which helps provide a more personalised and effective shopping experience. The social-network-based fashion recommendation system also leverages information about a user's social circles and fashion trends, further personalising recommendations and influencing purchasing decisions (Chakraborty et al., 2021).

Traditional recommender system approaches, such as collaborative and content-based filtering, have significant limitations when applied to fashion. These stem primarily from the irregularity of individual purchase histories and the inadequate representation of visual attributes, as the aesthetic complexity of fashion items is rarely captured by category names and metadata (Liao et al., 2022). Consequently, more recent research has shifted towards multimodal recommendation models that integrate richer sources of information. These models incorporate product images to convey colour, silhouette, and pattern; textual descriptions and customer reviews to communicate perceived fit and quality; and video-based content to capture movement and real-world garment presentation (Deldjoo et al., 2023).

Another challenge is predicting fit, which is a key driver of consumer dissatisfaction and product returns in online fashion retail. AI-enhanced systems now address this issue by using biometric data, purchase feedback and machine learning techniques to improve size and fit recommendations (Deldjoo et al., 2023). Therefore, these developments represent a significant shift towards models that are visually and contextually aware, rather than purely data-driven. However, as recommendation systems become increasingly adaptive and AI-driven, they also gain the capacity to influence consumer decisions in subtle and often unrecognised ways.

2.2. Manipulative Design and "Dark Patterns"

Manipulation refers to the intentional attempt to influence someone’s decisions or behaviour in a way that undermines their rational thought or ability to make their own decisions. (Buss, 1987) In psychology field, manipulation is a mind control, specifically abuse of power, which is domination. (Van Dijk, 2006) It’s different from persuasion, which allows space for informed consent. Manipulation often targets cognitive biases, emotional vulnerabilities, or informational asymmetries to steer choices in a concealed or coercive manner. According to Susser, Roessler and Nissenbaum (2019), manipulation in digital environments occurs when systems 'covertly influence users by targeting their subconscious vulnerabilities', often through opaque algorithms and behavioural data mining. In the context of digital commerce, this can include encouraging consumers to make purchases by using urgency cues, biased defaults or aesthetic triggers that override reflective decision-making. Although such strategies may boost engagement and conversions, they raise significant ethical issues regarding user self-determination and fairness (Mathur et al., 2019).

Manipulative design through the lens of "dark patterns”, which is in user experiences, a persuasive design that intentionally coerce and deceive users into actions against their best interests, such as nagging, obstruction and interface interference (Gary et al., 2018). Dark patterns are believed to be effective because they manipulate cognitive biases and human bounded rationality (Bösch et al., 2016; Waldman, 2020). AI-powered recommendation systems push fast fashion by continuously feeding users’ decisions tailored to preferences and emerging trends. This personalisation can lead to compulsive buying, creating endless loops of new suggestions that fuel overconsumption and environmental harm.

2.3 Manipulative Design in Fashion Recommendation Systems

AI-powered recommendation systems in the fashion industry not only predict consumer preferences, but also increasingly shape them. Through interface design and algorithmic personalisation, these systems are blurring the boundary between assistance and manipulation (Susser et al., 2019). In the context of fashion, an industry deeply linked to self-expression and identity, such influence takes on an emotional dimension, enhancing desires, insecurities, and social aspirations (Entwistle, 2015; Belk, 2013).

Urgency cues, such as “only 2 left in stock” or “sale ends in 3 hours” are among the most common dark patterns in online fashion platforms. These features increase consumers ’fear of missing out (FOMO) and loss aversion, psychological biases that trigger impulsive purchasing (Gray et al., 2018; Mathur et al., 2019). In algorithmic systems, urgency cues are often personalised, generated in real time based on user engagement data, thereby enhancing their persuasive power. While initially intended to drive efficiency in decision-making, these cues manipulate affective states like anxiety and excitement, resulting in reduced reflective thought before purchase.

Another dominant manipulative feature is social validation. Recommender systems frequently display messages such as '10 people bought this in the last hour' or 'popular among people like you'. These design patterns simulate social proof, targeting users' desire for acceptance and belonging to encourage conformity (Cialdini, 2009). On fashion platforms in particular, this social mirroring is powerfully effective because fashion choices are deeply social and identity-affirming (McCracken, 1986). Even when generated by algorithms, the perceived approval of others can alter individual taste and reinforce collective trends, aligning consumer behaviour more closely with commercial objectives than personal expression.

Aesthetic persuasion is more sophisticated yet equally powerful. AI-driven recommendation feeds are visually tailored to stimulate aesthetic desire through lighting, body imagery and visual repetition (Nasser, 2023). Platforms such as Instagram and Pinterest employ algorithmic aesthetics to create an illusion of harmony and belonging by matching colour tones, models, and visual style (Li et al., 2025). This aesthetic uniformity influences emotional engagement and perceived self-congruence, prompting users to equate algorithmic beauty with their own aspirations. Under these circumstances, the recommender system serves not only as a sales tool, but also as a cultural script that guides users in what to desire and how to construct their identity.

Research on this topic has so far focused on the efficiency of algorithms and data ethics, with few studies exploring the emotional impact of algorithmic manipulation on consumers. In the context of fashion, this emotional entanglement is compounded because the consumption of clothing interacts with body image, self-esteem, and the formation of identity (Dittmar, 2008). When users are exposed to aesthetically persuasive and socially validated recommendations, they may internalise these cues as aspirational signals, thereby reinforcing algorithmically mediated standards of beauty and taste. Therefore, algorithmic manipulation in fashion is an affective and ethical issue as well as a commercial one, as it changes how consumers feel, desire, and define themselves in digital retail spaces.

2.4 Ethical Considerations and Algorithmic Exploitation of Vulnerabilities

The ethical implications of AI-based recommendation systems extend beyond just user privacy and data security. The literature identifies more insidious forms of harm, particularly the potential for algorithms to exploit human vulnerabilities. Strümke et al. (2023) introduce the critical concept of machine learning models that imitate human characteristics which are considered vulnerabilities, such as young age, depression, or a tendency towards gambling addiction. This approach raises ethical alarms around emotional manipulation and mental health impact. The authors caution that decisions made by algorithms based on these 'vulnerable state representations' could be perceived as manipulative and exacerbate the situation of vulnerable individuals.

The fashion industry is a domain where exploitation is particularly concerning, given its emphasis on social trends, self-image, and impulse purchases. An AI system that identifies a user's insecurities or tendency to shop emotionally could, intentionally or not, make recommendations that reinforce these vulnerabilities. Previous scholars indicate that most participants were aware of manipulative design patterns but still felt they could not avoid or resist them in practice (Bongard-Blanchy et al., 2021), while some participants described ‘feeling manipulated even when they knew it’, expressing frustration and resignation toward online persuasive tactics. This tension, described by Elliott 2024, as a form of algorithmic anxiety, reflects a deeper uncertainty in which consumers begin to question whether their desires are genuinely their own or the product of algorithmic construction. As a result, participants expressed a desire for greater transparency from online platforms, seeking clearer communication about the intent behind personalised and persuasive design. Masciari et al. (2024)

emphasise the importance of designing fair and transparent systems to address these ethical implications, highlighting that responsible AI development is essential for achieving equitable and non-exploitative outcomes.

2.5 Research gaps and research question

Although AI-based recommendation systems have received extensive scholarly attention, most studies in the fashion domain remain technically driven, emphasising algorithmic design, prediction accuracy and system typology rather than user experience (Deldjoo et al., 2023; Shirkhani et al., 2023; Guan et al., 2025). This reveals a significant gap between the technological optimisation of recommendation systems and their psychological and ethical implications in fashion retail.

RS 1. How do AI-powered fashion recommendation systems manipulate consumer behaviour through design and algorithmic personalisation, and what are their perceptions and emotional responses?

Research on dark patterns has begun to address user experience more broadly. Bongard-Blanchy et al. (2021) called for further investigation into how users perceive manipulative design features, while Gray et al. (2021) found that users experience frustration, resignation and loss of control when confronted with manipulative interfaces. Liao et al. (2022) further explored how different algorithmic models affect user trust and credibility. However, none of these studies have examined these dynamics specifically in fashion, where visual, social and identity factors heighten emotional vulnerability.

RS 2. What ethical issues arise from the use of manipulative AI design in fashion recommender systems?

Masciari et al. (2024) identify manipulation and autonomy as major ethical challenges in AI-based recommender systems, yet acknowledge these dimensions remain largely conceptual rather than empirically tested. The fashion industry is closely tied to self-presentation, body image and social validation, that has notably been absent from this ethical discourse. This study addresses these gaps by empirically examining how manipulative design features in AI-powered fashion recommendation systems influence consumers' perceptions, emotional experiences, and sense of autonomy, contributing to a more ethically responsible approach to algorithmic innovation in digital fashion.

3 Methodology

This study adopts a qualitative research design using semi-structured interviews to gain in-depth insights into how consumers perceive and emotionally respond to manipulative features embedded in AI-powered fashion recommendation systems. Qualitative methods are particularly suited to capturing the subjective, emotional, and interpretive dimensions of consumer–technology interactions (Creswell, 2013; Denzin & Lincoln, 2018). Through open-ended dialogue, this approach enables a deeper understanding of how algorithmic persuasion and digital personalisation shape shopping behaviour and perceptions of autonomy in online fashion contexts.

A purposive sampling strategy was employed to recruit participants who represent active Gen Z fashion consumers in the United Kingdom. Generation Z, typically defined as individuals born between 1995 and 2010, are the most digitally immersed consumers, known for frequent online shopping and high engagement with social media and fashion trends. Six participants aged 20 to 30 were interviewed, all residing in the UK. This age group aligns with prior research identifying Gen Z as the dominant demographic shaping digital retail and algorithmic consumption.

Interviews were conducted in person between September and October 2025, which are essential when exploring affective responses to digital experiences. Each interview lasted approximately 20–30 minutes and followed a semi-structured format with open-ended questions concerning participants' experiences with personalised fashion recommendations, targeted advertising, and perceptions of manipulation or autonomy. All participants were informed of the study's purpose and provided consent prior to participation. Interviews were audio-recorded and transcribed verbatim for analysis. Data were analysed thematically following Braun and Clarke's (2006) six-phase framework. Initial codes were generated from the interview transcripts and then refined into broader themes that reflect consumer emotions, perceptions of AI-driven recommendation systems, and experiences of algorithmic persuasion.

4 Findings

The thematic analysis revealed the following four main findings: (1) AI-powered fashion recommendation systems are perceived as both helpful and manipulative; (2) the most salient manipulative mechanisms are urgency, repetition and social proof; (3) participants are aware of these tactics yet still experience emotional pressure; and (4) there are ethical concerns around autonomy, privacy and creativity. Overall, the findings

demonstrate that fashion recommendation systems not only assist with product discovery, but also mediate tacit, emotional and relational forms of knowledge. This shapes what consumers find desirable and how they interpret their own preferences.

4.1 Between convenience and coercion: recommendation as both helpful and manipulative

Participants rarely viewed recommendation systems as either entirely positive or entirely negative, describing them as simultaneously useful and manipulative. Many appreciated personalised suggestions that saved time and matched their taste or browsing behaviour. Participant 5 noted that recommendations were useful because they "speed up finding what you want." However, this usefulness was consistently qualified by discomfort. Participant 6 argued that recommendation systems ""puts you in a little box you never come out of... They're always going to go for what you're always on. But I'm not always on the same thing, my style changes all the time. I feel like the recommendation ruins it for me." It reflects that AI recommendation system effectively reduces individuality and encourages consumers to "dress like everyone else." Participant 4 reflected that repeated exposure can create a false sense of desire, making consumers think "it's my own idea" when a product has simply been pushed repeatedly. Participants did not reject personalisation itself, rather, they objected to assistance becoming coercive.

4.2 Manipulative mechanisms: repetition, scarcity and social validation

Three mechanisms emerged as especially prominent: repetition, urgency and social proof. Repetition was among the most frequently cited tactics, with participants reporting exposure to the same products across multiple platforms after a single search. Participant 5 recalled an advertisement appearing "over 20 times" even after clicking "not interested", while Participant 4 described cross-platform tracking from Google to Instagram as "weird" and subtly influential.

Scarcity and urgency cues, such as "discount ends in 24 hours" or "only two left in stock" were widely recognised as manipulative yet acknowledged as effective. Participants admitted that countdown timers could influence their behaviour, particularly when prices genuinely increased afterwards. As Participant 6 admitted, "Even though I know it's a marketing trick, I still sometimes fall for it." These responses highlight that scarcity cues are effective not because consumers fail to recognise them, but because they trigger emotional states that interfere with rational decision-making.

Social validation messages such as "545 people bought this" produced more mixed reactions. Participant 3 stated that if too many people had purchased the same item, she would be less inclined to want it, valuing individuality over conformity. Even this rejection confirms the role of relational knowledge in fashion consumption, with participants interpreting such prompts through the lens of social comparison, belonging and distinction.

4.3 Awareness without resistance: emotional responses to algorithmic persuasion

A key finding was that awareness of manipulation did not prevent participants from experiencing its emotional effects. Most understood how algorithms used browsing history and behavioural data to shape recommendations, yet this awareness coexisted with frustration, resignation and unease. Participant 3 described her phone as "stalking" her, while Participant 5 openly acknowledged purchasing products she did not need following repeated algorithmic exposure.

Participants were not simply misled by false information, but nudged emotionally through anxiety, urgency, and repetition. This indicates that algorithmic persuasion not only alters what consumers know, but also shapes how they feel about their own decision-making.

4.4 Autonomy, privacy and ethical concerns

Most participants claimed a sense of control over their purchases, yet often framed this autonomy as something requiring active defence through planning ahead, rejecting cookies, or comparing options across multiple sites. Even those who felt personally resistant expressed concern that less experienced shoppers might be more vulnerable. Privacy concerns were significant, with participants uncomfortable about the extent of data collection even when accepting it as unavoidable. Participant 3 felt platforms wanted "height, weight, blood... literally everything", while Participant 4 questioned why brands should use his data when he could simply search for products himself. Some were willing to tolerate data use if it improved relevance, indicating an uneasy trade-off between convenience and surveillance. These responses emphasise the fundamental ethical conflict

between personalisation and the right to informed consent. Users are rarely given meaningful control over how their data is collected, processed, or used to influence their behaviour.

Participants also expressed concern that recommendation systems were limiting their sense of self and creative freedom beyond individual purchases. Participant 3 argued that AI "strips the artist from the artwork", while Participant 6 suggested that excessive recommendations made it "hard to hear your own opinion", particularly significant in fashion, where consumption is closely tied to identity and self-expression. These concerns are particularly significant in the context of fashion, where consumption is closely tied to identity formation and self-expression. When recommendation systems encourage conformity and restrict exploration, they raise ethical questions about both individual autonomy and the cultural implications of algorithmically mediated style.

Despite their varying levels of technical awareness, participants in the study all expressed a desire for greater transparency regarding how recommendations are generated and how personal data is used. Participant 2 stated that 'too many recommendations annoy customers, transparency is better', while Participant 4 noted that some platforms are 'manipulative and it's unclear how they use your data', suggested that transparency should at least be offered as an option. Some people might find it scary or intrusive, especially older adults who aren't familiar with social media, more transparency could be more ethical. These perspectives reveal a clear consumer expectation for greater accountability from platforms, demonstrating that the current opacity of algorithmic systems represents an ethical deficit in the digital fashion industry.

Overall, the findings show that AI-powered fashion recommendation systems are perceived as providing both assistance and influence. Although participants valued personalisation that was relevant to them, they became critical when systems relied on repetition, scarcity and social validation to influence behaviour. Thus, awareness of manipulation did not only eliminate its emotional impact, and the accounts of participants reveal a variety of ethical concerns, ranging from the covert use of data and deceptive design to the erosion of creative autonomy.

5 Conclusion

This paper explores the intersection between AI-powered fashion recommendation systems and the emotionally manipulative design practices embedded within them. As these systems evolve and become more sophisticated by incorporating deep learning, multimodal data and affective profiling, they transform from passive tools of convenience into active shapers of consumer identity, self-perception and purchasing behaviour. The findings argue that algorithmic fashion curation is not neutral, instead, it draws on tacit emotional knowledge and exploits well-documented cognitive biases, employing dark patterns that undermine individual autonomy.

The fashion industry is particularly vulnerable to manipulation. Due to its close links with body image, social validation and self-expression, consumers are susceptible to personalisation systems based on insecurities rather than genuine preferences. The convergence of scarcity cues, social proof and identity-targeted recommendations in a single interface creates a situation where desire is manufactured, blurring the line between autonomous choice and algorithmic influence.

Therefore, policymakers must extend their oversight to include affective manipulation as a distinct category of harm, going beyond transparency requirements such as the disclosure of algorithmic decision-making process. Regulatory bodies should consider mandating 'dark pattern audits' for AI-powered retail platforms, paying particular attention to features that exploit emotional vulnerability or body-related insecurities. Those engaged in the realm of fashion AI, encompassing designers and developers, bear an ethical responsibility to distinguish between personalisation that assists the user and that which prioritises conversion metrics at the expense of the user's psychological well-being. It is crucial to adopt value-sensitive design principles.

Consumers should be aware that simply being aware of manipulative design is not enough to protect against its effects. However, digital literacy initiatives that help consumers recognise emotional targeting, artificial urgency, and algorithmic self-image construction can meaningfully shift the power balance. Educating users about how their emotional data is inferred and used is a necessary step towards more informed and autonomous consumption.

This paper highlights a gap in the literature to be addressed in future research: the psychological and emotional dimensions of AI-driven fashion recommendations are comparatively under-explored in relation to technical performance metrics. Future research could focus on longitudinal studies examining the effects of prolonged exposure to emotionally targeted fashion AI on body image, self-esteem, and purchasing autonomy, particularly among younger and more vulnerable demographic groups. Qualitative approaches based on the

consumers' lived experiences will help to capture the tacit, subjective dimensions of algorithmic influence, that quantitative methods struggle to reach.

6 Ethics Declaration

This research involved semi-structured interviews with human participants, followed by established research ethics standards. All participants were fully informed of the study's purpose, their right to withdraw at any time without consequence, and how their data would be used and stored. Written informed consent was secured from each participant before data collection commenced. The principles of anonymity, confidentiality, and data protection were strictly adhered to. No conflicts of interest are declared.

7 AI Declaration

No AI tools were used in the conceptualisation, research design, data preparation, data analysis, presentation of findings, and list of references. AI language tool ChatGPT by OpenAI was used selectively during the revision and editing stage of this paper to support grammar checking, sentence refinement, and proofreading. The authors carefully reviewed and verified AI-assisted output before inclusion in the final manuscript. All citations and references were checked by the authors. The authors acknowledge the limitations of AI tools, including the potential for bias, errors, and gaps in knowledge.

Reference

Bongard-Blanchy, K., Rossi, A., Rivas, S., Sur Alzette, E., Doublet, S., Koenig, V. and Lenzini, G. (2021). 'I am Definitely Manipulated, Even When I am Aware of it. It's Ridiculous! '-Dark Patterns from the End-User Perspective. *"I am Definitely Manipulated, Even When I am Aware of it. It's Ridiculous!" - Dark Patterns from the End-User Perspective*. doi:https://doi.org/10.1145/3461778.3462086.

Bösch, C., Erb, B., Kargl, F., Kopp, H. and Pfattheicher, S. (2016). Tales from the Dark Side: Privacy Dark Strategies and Privacy Dark Patterns. *Proceedings on Privacy Enhancing Technologies*, 2016(4), pp.237–254. doi:https://doi.org/10.1515/popets-2016-0038.

Buss, D. M. (1987). Selection, evocation, and manipulation. *Journal of Personality and Social Psychology*, *53*(6), 1214–1221. https://doi.org/10.1037/0022-3514.53.6.1214

Chakraborty, S., Hoque, M.S., Rahman Jeem, N., Biswas, M.C., Bardhan, D. and Lobaton, E. (2021). Fashion Recommendation Systems, Models and Methods: A Review. *Informatics*, 8(3), p.49. doi:https://doi.org/10.3390/informatics8030049.

Deldjoo, Y., Nazary, F., Ramisa, A., McAuley, J., Pellegrini, G., Bellogin, A. and Noia, T.D. (2024). A Review of Modern Fashion Recommender Systems. *ACM Computing Survey*, 56(4), pp.1–37. doi:https://doi.org/10.1145/3624733.

Dittmar, H. (2008). *Consumer Culture, Identity and Well-Being*. Psychology Press. doi:https://doi.org/10.4324/9780203496305.

Elliott, A. (2024). *Algorithms of Anxiety*. John Wiley & Sons.

Fogg, B. (2009). A Behavior Model for Persuasive Design. *Proceedings of the 4th International Conference on Persuasive Technology - Persuasive '09*, 40(40), pp.1–7. doi:https://doi.org/10.1145/1541948.1541999.

Geetha, M.P. and Renuka, D.Karthika. (2019). Research on Recommendation Systems using Deep Learning Models. *International Journal of Recent Technology and Engineering*, 8(4), pp.10544–10551. doi:https://doi.org/10.35940/ijrte.d4609.118419.

Gray, C.M., Chen, J., Chivukula, S.S. and Qu, L. (2021). End User Accounts of Dark Patterns as Felt Manipulation. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW2), pp.1–25. doi:https://doi.org/10.1145/3479516.

Gray, C.M., Kou, Y., Battles, B., Hoggatt, J. and Toombs, A.L. (2018). The Dark (Patterns) Side of UX Design. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*, pp.1–14. doi:https://doi.org/10.1145/3173574.3174108.

Guan, C., Qin, S., Ling, W. and Ding, G. (2016). Apparel recommendation system evolution: an empirical review. *International Journal of Clothing Science and Technology*, 28(6), pp.854–879. doi:https://doi.org/10.1108/ijcst-09-2015-0100.

Li, H.-C., Wang, L.-K., Chang, Y.-K. and Huang, K.-Y. (2025). Establishing colour harmony evaluation and recommendation model for clothing colour matching based on machine learning and deep learning. *Fashion and Textiles*, 12(1). doi:https://doi.org/10.1186/s40691-025-00433-y.

Liao, M., Sundar, S.S. and B. Walther, J. (2022). User Trust in Recommendation Systems: A comparison of Content-Based, Collaborative and Demographic Filtering. *CHI Conference on Human Factors in Computing Systems*. doi:https://doi.org/10.1145/3491102.3501936.

Masciari, E., Umair, A. and Ullah, M.H. (2024). A Systematic Literature Review on AI-Based Recommendation Systems and Their Ethical Considerations. *IEEE Access*, 12, pp.121223–121241. doi:https://doi.org/10.1109/access.2024.3451054.

Mathur, A., Acar, G., Friedman, M.J., Lucherini, E., Mayer, J., Chetty, M. and Narayanan, A. (2019). Dark Patterns at Scale: Findings from a Crawl of 11K Shopping Websites. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW), pp.1–32. doi:https://doi.org/10.1145/3359183.

McCracken, G. (1986). Culture and Consumption: a Theoretical Account of the Structure and Movement of the Cultural Meaning of Consumer Goods. *Journal of Consumer Research*, 13(1), pp.71–84.

Nasser, A. (2023). Applications of Artificial Intelligence in the field of media. *International Journal of Artificial Intelligence and Emerging Technology*, 6(2), pp.19–41. doi:https://doi.org/10.21608/ijaiet.2024.275179.1006.

Resnik, A.J. and Cialdini, R.B. (1986). Influence: Science & Practice. *Journal of Marketing Research*, 23(3), p.305. doi:https://doi.org/10.2307/3151490.

Shirkhani, S., Mokayed, H., Saini, R. and Chai, H.Y. (2023). Study of AI-Driven Fashion Recommender Systems. *SN Computer Science*, [online] 4(5). doi:https://doi.org/10.1007/s42979-023-01932-9.

Strümke, I., Marija Slavkovik and Clemens Stachl (2023). Against Algorithmic Exploitation of Human Vulnerabilities. *arXiv (Cornell University)*. doi:https://doi.org/10.48550/arxiv.2301.04993.

Susser, D., Roessler, B. and Nissenbaum, H. (2019). Technology, autonomy, and manipulation. *Internet Policy Review*, 8(2). doi:https://doi.org/10.14763/2019.2.1410.

Van Dijk, T. A. (2006). Discourse and manipulation. *Discourse & Society*, *17*(3), 359-383. https://doi.org/10.1177/0957926506060250

Waldman, A.E. (2020). Cognitive biases, dark patterns, and the 'privacy paradox'. *Current Opinion in Psychology*, 31(1), pp.105–109. doi:https://doi.org/10.1016/j.copsyc.2019.08.025.

Zarouali, B., Boerman, S.C., Voorveld, H.A.M. and van Noort, G. (2022). The algorithmic persuasion framework in online communication: conceptualization and a future research agenda. *Internet Research*, ahead-of-print(ahead-of-print). doi:https://doi.org/10.1108/intr-01-2021-0049.

Project Tailoring as a Knowledge Legitimacy Process in Hybrid Intelligence Environments

Anikó Csepregi[1], Coşkun Akça[2]
[1]University of Pannonia, Veszprém, Hungary
[2]Kastamonu University, Kastamonu, Türkiye

Abstract

Organizational change is increasingly implemented through projects that operate as temporary organizational settings where actors interpret uncertainty, negotiate priorities, and enact change through coordinated action. While project tailoring is traditionally understood as a technical adaptation of project management methods, recent research suggests that tailoring decisions emerge through collective sensemaking processes embedded in organizational and technological contexts. However, existing literature has paid limited attention to how organizational culture and digital technologies jointly shape knowledge processes through which tailoring decisions are enacted.

This paper explains how organizational culture and hybrid intelligence jointly shape knowledge legitimacy processes through which decisions are enacted in project-based organizations. This conceptual paper develops an integrative framework explaining project tailoring as a knowledge enactment process emerging from collective sensemaking and knowledge legitimacy processes. More specifically, the paper conceptualizes project tailoring as a knowledge legitimacy process emerging from collective sensemaking in hybrid intelligence environments. The framework integrates literature on project-based organizing, sensemaking, organizational culture, and sociomateriality, and introduces hybrid intelligence as a socio-technical context that shapes the visibility, measurability, interpretation, and legitimacy of knowledge in project environments.

The paper argues that project tailoring decisions should be understood as outcomes of culturally and sociomaterially embedded knowledge processes rather than purely technical adjustments. Organizational culture is conceptualized as an interpretive and knowledge logic that shapes how actors evaluate and legitimize knowledge, while hybrid intelligence environments influence what knowledge becomes visible, measurable, and actionable in decision-making processes. Project tailoring is therefore interpreted as an enacted outcome of interacting cultural, interpretive, and technological processes.

The paper contributes to knowledge management and project studies by reframing project tailoring as a culturally embedded knowledge process, introducing knowledge legitimacy as a key mechanism linking collective sensemaking and project decision making, and by conceptualizing hybrid intelligence as a sociomaterial context influencing organizational knowledge processes and decision outcomes. The paper concludes by outlining implications for organizational change and knowledge-based decision making in hybrid intelligence environments.

Keywords: organizational culture; hybrid intelligence; knowledge legitimacy; collective sensemaking; project-based organizations

Paper type: Academic Research Paper

1 Introduction

Organizational change is frequently enacted through projects that function as temporary organizational settings in which new structures, practices, and meanings are negotiated under conditions of uncertainty (Packendorff, 1995; Cicmil et al., 2006; Sloot et al., 2024; Antonova, 2025). In such contexts, projects are not merely execution mechanisms but organizational arenas where actors interpret situations, negotiate priorities, and enact change through coordinated action (de Metz et al., 2024). Project tailoring therefore cannot be understood only as a technical adjustment to project characteristics, but also as a socially embedded process through which actors interpret situations, negotiate priorities, and justify decisions in uncertain environments (Söderlund, 2011; Lundin and Söderholm, 1995; Sydow and Windeler, 2020).

Within project-based organizing, project tailoring has traditionally been understood as a mechanism for adapting project management practices to contextual conditions such as project size, complexity, or uncertainty (Miterev et al., 2017; Araki and Martins, 2023). This contingency-based perspective conceptualizes tailoring as a rational adjustment process aimed at achieving fit between project characteristics and management approaches. However, this view under-theorizes the interpretive and social processes through which adjustments occur (Whittle et al., 2023). Project tailoring can therefore be understood not only as a technical adjustment but also as a socially embedded process of interpretation and negotiation, highlighting the need to examine how such decisions are socially constructed and justified in project environments (Söderlund, 2011; Lundin and Söderholm, 1995; Sydow and Windeler, 2020).

Research on sensemaking shows that decisions in uncertain organizational environments emerge through collective processes of interpretation through which actors construct shared meanings and justify action (Maitlis and Christianson, 2014; Whittle et al., 2023). These processes are shaped by organizational culture, which shapes how situations are interpreted and what forms of knowledge are considered legitimate for decision making (Schein, 2010; Pagnozzi et al., 2024). At the same time, digital technologies and artificial intelligence increasingly shape organizational decision making by influencing knowledge visibility, interpretation, and coordination, highlighting the sociomaterial nature of contemporary organizational environments (Orlikowski, 2007; Ledro et al., 2023; Sherson, 2024). Hybrid intelligence environments therefore not only support decision-making but also shape how knowledge is produced, evaluated, and applied in organizational contexts (Shrestha et al., 2019; Sherson, 2024).

Despite growing research on project tailoring, organizational culture, sensemaking, and hybrid intelligence, these streams have largely been examined separately. As a result, we know relatively little about how project tailoring decisions emerge as outcomes of collective sensemaking and knowledge legitimacy processes embedded in cultural and technological contexts. This fragmentation indicates the need for an integrative perspective that connects project studies, sensemaking theory, knowledge management, and sociomateriality in order to better understand project-level decision making in contemporary organizational environments (Sydow and Windeler, 2020; Sandberg and Tsoukas, 2015).

This paper addresses this gap by explaining how organizational culture and hybrid intelligence jointly shape knowledge legitimacy processes through which decisions are enacted in project-based organizations. By doing so, the paper shifts the focus from project tailoring as a technical adjustment to project tailoring as a knowledge-based and socially constructed organizational process and develops a process perspective on project decision making based on collective sensemaking and knowledge legitimacy (Maitlis and Christianson, 2014; Sandberg and Tsoukas, 2015; Andrews et al., 2024; Sydow and Windeler, 2020).

This paper positions project tailoring within process and knowledge-based perspectives on organizational decision making. Rather than conceptualizing tailoring as a contingency-based adaptation or methodological choice, the paper conceptualizes project tailoring as an outcome of knowledge legitimacy processes through which certain interpretations become accepted as valid bases for action. This shifts the focus from project methodologies to the social and sociomaterial processes through which decisions become justified and enacted in project environments. In this sense, the paper contributes to project studies by explaining project tailoring as a knowledge selection process, to sensemaking research by linking interpretation to decision outcomes through legitimacy, and to sociomaterial research by conceptualizing hybrid intelligence as a cultural actor shaping what knowledge becomes legitimate in organizational decision making.

The research question guiding this paper is:

How do organizational culture and hybrid intelligence configurations co-shape collective sensemaking and knowledge legitimacy processes through which project tailoring decisions are enacted during organizational change?

This paper makes three theoretical contributions. First, it contributes to project studies by reconceptualizing project tailoring as an outcome of knowledge legitimacy processes rather than a technical adaptation task. Second, it contributes to sensemaking and knowledge management literature by introducing knowledge legitimacy as a knowledge validation and selection mechanism linking collective sensemaking and project decisions. Third, it contributes to sociomaterial and hybrid intelligence research by explaining how digital technologies and AI systems shape knowledge visibility, interpretation, and legitimacy in project environments.

The remainder of the paper is structured as follows. The next section reviews the theoretical background, followed by the development of the conceptual framework. The paper then discusses theoretical contributions and implications, and concludes with limitations and directions for future research.

2 Theoretical Background

2.1 Projects and Organizational Change

In organizations, change is often implemented through projects that function as temporary settings in which actors interpret uncertainty, negotiate priorities, and make decisions (Lundin and Söderholm, 1995; Packendorff, 1995). Projects are therefore not merely execution units but organizational spaces where change is interpreted and enacted through collective processes (Sydow and Windeler, 2020; Sloot et al., 2024). Because project environments are characterized by uncertainty and evolving conditions, decisions cannot be fully predefined, and actors must continuously interpret situations and adapt their actions (Söderlund, 2011; Clegg et al., 2018).

In this context, project tailoring has traditionally been understood as a contingency-based alignment process aimed at achieving fit between project management methods and project characteristics (Miterev et al., 2017; Araki and Martins, 2023). However, recent research suggests that tailoring decisions are not purely technical adjustments but the result of interactions, interpretations, and negotiations between competing knowledge claims (de Metz et al., 2024; Brill et al., 2025). Project tailoring should therefore be understood as a socially embedded decision in which certain interpretations become accepted as legitimate bases for action in a given organizational context (Abimbola et al., 2024; Andrews et al., 2024).

2.2 Sensemaking and knowledge legitimacy

In uncertain and complex organizational environments, decisions emerge through collective sensemaking processes in which organizational actors jointly interpret situations and construct shared meanings that guide action (Maitlis and Christianson, 2014; Sandberg and Tsoukas, 2015). However, not all interpretations become the basis for decisions. Decision making is shaped by knowledge that organizational actors accept as legitimate, relevant, and credible (Deephouse et al., 2017; Andrews et al., 2024). Knowledge legitimacy is therefore the outcome of social processes in which different knowledge claims compete and some become accepted as valid bases for action while others remain marginal (Abimbola et al., 2024).

Organizational decisions can thus be understood as outcomes of two linked processes: collective sensemaking, which produces interpretations, and knowledge legitimacy, which selects which interpretations become accepted as the basis for decisions.

2.3 Organizational culture as interpretive and knowledge logic

Organizational culture can be understood as an interpretive system that shapes how organizational actors interpret situations and what knowledge they consider relevant and acceptable for decision making (Harris, 1994; Long and Fahey, 2000). In this sense, culture not only provides values and norms but also acts as a knowledge evaluation logic that determines what knowledge becomes legitimate for decision making (Adeinat and Abdulfatah, 2019).

This is particularly important in project environments, where temporary organizational forms rely heavily on interpretive processes and where different professional groups possess different forms of knowledge (Clegg et al., 2018; Wei and Miraglia, 2017). Organizational culture therefore shapes which knowledge sources are considered credible and relevant and what decision logic guides decision making (Pagnozzi et al., 2024; Pescatore et al., 2022). Organizational decision making can thus be understood as a culturally embedded process in which culture shapes both sensemaking and knowledge legitimacy.

2.4 Sociomateriality, Technology, Hybrid Intelligence

The sociomaterial perspective suggests that technology and organizational practices are intertwined in shaping organizational functioning and decision making (Orlikowski, 2007). Digital technologies and artificial intelligence increase the visibility and measurability of organizational processes, thereby influencing what information and forms of knowledge become the basis of decisions (Olan et al., 2022; Ledro et al., 2023). Technology therefore does not merely support decision making but shapes what becomes visible, measurable, and legitimate knowledge in organizations (Andrews et al., 2024; Brill et al., 2025).

With the emergence of artificial intelligence, decision making increasingly takes place in hybrid intelligence systems where human and artificial intelligence jointly contribute to decision processes (Jarrahi, 2018; Shrestha et al., 2019; Sherson, 2024). Decisions thus emerge in sociomaterial practices where human interpretation and algorithmic analysis jointly shape the knowledge that becomes the basis for decisions (Fabri et al., 2023; Orlikowski, 2007).

2.5 Integrative theoretical framework

This paper integrates project studies, sensemaking theory, knowledge management, and sociomateriality in order to explain how project tailoring decisions emerge in contemporary organizations. Project studies conceptualize projects as temporary organizational contexts in which decisions are enacted under conditions of uncertainty (Lundin and Söderholm, 1995; Sydow and Windeler, 2020). Sensemaking theory explains how actors interpret uncertain situations and construct shared meanings (Maitlis and Christianson, 2014; Sandberg and Tsoukas, 2015), while knowledge management research explains how certain knowledge becomes validated and accepted as a basis for decisions (Deephouse et al., 2017; Abimbola et al., 2024). Sociomateriality and hybrid

intelligence further explain how digital technologies and AI systems shape the visibility and measurability of knowledge in organizational decision-making (Orlikowski, 2007; Shrestha et al., 2019; Sherson, 2024).

By linking these perspectives, the paper develops a process framework in which collective sensemaking produces interpretations, knowledge legitimacy selects which interpretations become accepted as valid knowledge, and these legitimized knowledge claims form the basis of project tailoring decisions. Organizational culture shapes what knowledge is considered legitimate, while hybrid intelligence shapes what knowledge becomes visible and measurable. Project tailoring is therefore conceptualized as the outcome of a culturally and technologically embedded process of sensemaking and knowledge legitimacy.

3 Conceptual Framework

This paper develops a conceptual framework that explains project tailoring as a culturally embedded knowledge enactment process emerging from collective sensemaking and knowledge legitimacy processes in hybrid intelligence environments. By integrating project-based organizing, sensemaking, organizational culture, and sociomaterial perspectives, the framework explains how project tailoring decisions emerge through collective interpretation and knowledge legitimation rather than purely technical decision making.

This paper argues that project tailoring should be understood not as a purely technical adjustment process but as a knowledge enactment process shaped by organizational culture, collective sensemaking, knowledge legitimacy, and hybrid intelligence.

3.1 Organizational culture as interpretive and knowledge logic

The framework conceptualizes organizational culture as an interpretive infrastructure that shapes how actors understand situations, evaluate knowledge, and justify decisions (Schein, 2010; Maitlis and Christianson, 2014). In project environments characterized by ambiguity and time pressure, actors rely on shared assumptions and culturally embedded expectations when interpreting situations and making decisions.

Organizational culture can therefore be understood as a knowledge logic that shapes what is considered credible and relevant knowledge in project decision making (Pagnozzi et al., 2024; Pescatore et al., 2022). Different cultural contexts privilege different forms of knowledge, such as formal methodologies, expert judgment, data-driven insights, or experiential knowledge. Culture thus shapes which knowledge becomes legitimate and is used in project tailoring decisions.

3.2 Collective sensemaking in project environments

Project decision making is conceptualized as a collective sensemaking process through which actors construct shared meanings and coordinate action in uncertain situations (Maitlis and Christianson, 2014; Whittle et al., 2023). During organizational change, project actors interpret situations, define problems, and negotiate appropriate courses of action. Project tailoring decisions therefore emerge through interaction and interpretation rather than purely analytical decision making.

Sensemaking also explains why project decisions are socially constructed and path-dependent, as decisions depend on dominant interpretations developed through interaction and organizational context.

3.3 Knowledge legitimacy as a validation and selection mechanism

A central concept in the framework is knowledge legitimacy, defined as the process through which certain knowledge claims become accepted as valid and appropriate for decision making (Andrews et al., 2024; Abimbola et al., 2024). Knowledge legitimacy functions as a validation and selection mechanism that determines which knowledge becomes actionable in project environments.

Organizational culture shapes evaluation criteria by influencing what is considered credible and relevant knowledge, while collective sensemaking shapes how actors interpret situations and evaluate competing knowledge claims. Knowledge legitimacy therefore explains how interpretations are translated into concrete project decisions and practices.

3.4 Hybrid intelligence as a sociomaterial context

The framework incorporates a sociomaterial perspective by recognizing that digital technologies and artificial intelligence systems influence knowledge processes and decision making (Orlikowski, 2007; Jansson et al., 2020). Technologies shape what becomes visible, measurable, and analyzable in organizations, thereby influencing knowledge evaluation and legitimacy.

From a hybrid intelligence perspective, decision making increasingly emerges from the interaction between human actors and technological systems (Sherson, 2024). Technology therefore acts as a sociomaterial context that shapes knowledge visibility, interpretation, and coordination processes in project environments.

3.5 Project tailoring as knowledge enactment

Based on these mechanisms, project tailoring is conceptualized as the outcome of culturally embedded collective sensemaking and knowledge legitimacy processes in hybrid intelligence environments. Project tailoring is therefore understood as a knowledge enactment process through which actors translate interpretations and legitimized knowledge into project structures, practices, and governance mechanisms.

The framework proposes that project tailoring emerges from the interaction of four elements: organizational culture, collective sensemaking, knowledge legitimacy, and hybrid intelligence. Organizational culture shapes interpretive frameworks, collective sensemaking constructs shared meanings, knowledge legitimacy determines which knowledge is accepted for decision making, and hybrid intelligence shapes knowledge visibility and interpretation. Together, these elements explain how project tailoring decisions are enacted during organizational change.

3.6 Conceptual model

Figure 1 presents the conceptual model developed in this paper. The model explains project tailoring as an outcome of culturally embedded collective sensemaking and knowledge legitimacy processes in hybrid intelligence environments. Organizational culture shapes interpretation and knowledge evaluation (Schein, 2010), collective sensemaking constructs shared understanding (Maitlis and Christianson, 2014; Whittle et al., 2023), knowledge legitimacy determines which knowledge becomes actionable (Andrews et al., 2024), and hybrid intelligence shapes knowledge visibility and evaluation (Orlikowski, 2007; Sherson, 2024).

The model proposes that knowledge legitimacy functions as a knowledge validation and selection mechanism that determines which knowledge claims become accepted and translated into project-level decisions and practices. Project tailoring is therefore conceptualized as the enacted outcome of culturally embedded collective sensemaking and knowledge legitimacy processes shaped by hybrid intelligence.

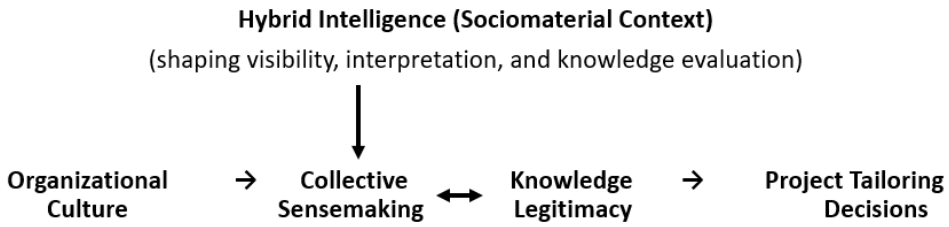


Figure 1: Conceptual model of project tailoring as a knowledge legitimacy process in hybrid intelligence environments.

The model does not conceptualize hybrid intelligence as a mediating variable but as a sociomaterial context that shapes knowledge visibility, interpretation, and evaluation throughout the decision process. Hybrid intelligence influences what becomes visible, measurable, and legitimate knowledge, thereby shaping both collective sensemaking and knowledge legitimacy processes through which project tailoring decisions are enacted.

The arrows in the model represent shaping relationships rather than linear causal effects. The model should therefore be interpreted as a process model that explains how project tailoring decisions emerge from interacting cultural, interpretive, and sociomaterial processes rather than as a variance-based causal model. Project tailoring is thus conceptualized as an emergent outcome of knowledge legitimacy processes embedded in organizational culture and hybrid intelligence environments.

4 Methodology

This paper is a theory-building conceptual study that develops a conceptual framework explaining how project tailoring decisions emerge in organizational change. Conceptual research focuses on developing new theoretical relationships and mechanisms rather than testing existing theories (Jaakkola, 2020; Whetten, 1989). The paper applies theory synthesis and conceptual integration by combining project studies, sensemaking theory, knowledge management, and sociomateriality to develop an integrative process model (MacInnis, 2011).

The aim is to explain the mechanisms through which project tailoring decisions emerge; therefore, the paper develops a process model linking collective sensemaking, knowledge legitimacy, and hybrid intelligence in project decision-making.

5 Discussion

This paper contributes to the literature by reframing project tailoring as a culturally embedded knowledge enactment process and by introducing knowledge legitimacy as a mechanism linking collective sensemaking and project-level decisions in hybrid intelligence environments. In doing so, it shifts the understanding of project tailoring from a contingency-based adjustment perspective toward a knowledge legitimacy perspective that explains how decisions emerge through socially and technologically embedded knowledge processes.

5.1 Project tailoring as knowledge enactment and knowledge legitimacy

This paper argues that project tailoring should be understood not primarily as a technical adjustment but as a knowledge enactment process emerging from collective interpretation and legitimation. While traditional literature conceptualizes tailoring as contingency-based adaptation (Miterev et al., 2017; Araki and Martins, 2023), this paper shows that tailoring decisions emerge through collective sensemaking processes in which actors interpret situations and justify action under uncertainty (Maitlis and Christianson, 2014; Whittle et al., 2023). By introducing knowledge legitimacy, the paper explains how certain interpretations become accepted as valid knowledge and translated into project practices, shifting the focus from analytical choice to culturally embedded knowledge validation and selection.

5.2 Hybrid intelligence as a sociomaterial context

The paper contributes to sociomateriality research by conceptualizing hybrid intelligence as a context that shapes knowledge visibility and evaluation. Technologies influence what becomes visible and measurable, thereby shaping which knowledge claims can become legitimate (Orlikowski, 2007; Jansson et al., 2020). From a hybrid intelligence perspective, decision-making emerges from the interaction between human and technological actors (Shrestha et al., 2019; Sherson, 2024). Technology thus actively participates in knowledge legitimacy processes by shaping what counts as credible knowledge and can be understood as a cultural actor co-shaping meaning and decision-making processes.

5.3 Theoretical contributions and managerial implications

The paper makes three main theoretical contributions. First, it reconceptualizes project tailoring as a knowledge enactment process rather than a methodological adjustment, extending project studies beyond contingency-based views (Miterev et al., 2017; Sydow and Windeler, 2020). Second, it introduces knowledge legitimacy as a mechanism linking collective sensemaking and project decision-making (Maitlis and Christianson, 2014; Andrews et al., 2024). Third, it conceptualizes hybrid intelligence as a sociomaterial context shaping knowledge visibility and legitimacy (Orlikowski, 2007; Shrestha et al., 2019).

From a managerial perspective, project tailoring should not be treated as a technical decision but as a knowledge legitimacy process in which competing knowledge claims are evaluated and made visible through collective sensemaking. Managers should move beyond selecting tools and actively govern how knowledge is validated and legitimized. This includes balancing multiple knowledge forms, ensuring transparency in decision criteria, and recognizing that organizational culture shapes what knowledge is accepted. In hybrid intelligence environments, managers must recognize that AI and digital tools shape knowledge visibility and thus influence decision outcomes, requiring the orchestration of interactions between human judgment and algorithmic reasoning.

5.4 Limitations and future research

The main limitation of this paper lies in its conceptual nature, as the proposed model has not been empirically tested. Future research could examine how collective sensemaking and knowledge legitimacy processes unfold in projects through qualitative, longitudinal, and process-oriented designs (Maitlis and Christianson, 2014; Orlikowski, 2007; Sherson, 2024). Multi-level studies could further explore how these dynamics evolve across individual, team, and organizational levels. Further research is also needed to investigate how hybrid intelligence reshapes knowledge legitimacy, particularly when algorithmic recommendations and human judgment diverge, and how organizational culture influences these processes in practice.

6 Conclusions

This paper set out to explain how organizational culture and hybrid intelligence jointly shape knowledge legitimacy processes through which decisions are enacted in project-based organizations. The paper showed that project tailoring decisions emerge from culturally and sociomaterially embedded processes of collective sensemaking and knowledge legitimation. Projects should therefore be understood as temporary organizational settings in which decisions emerge through collective interpretation and evaluation of knowledge rather than purely technical decision making (Packendorff, 1995; Sydow and Windeler, 2020; Maitlis and Christianson, 2014).

The paper developed a conceptual framework that explains project tailoring as a culturally embedded knowledge enactment process in which organizational culture shapes knowledge evaluation, collective sensemaking constructs interpretations, knowledge legitimacy determines which knowledge becomes accepted for decision making, and hybrid intelligence shapes knowledge visibility and measurability (Schein, 2010; Orlikowski, 2007; Andrews et al., 2024). Project tailoring is therefore conceptualized as an emergent outcome of interacting cultural, interpretive, and sociomaterial processes.

The main contribution of the paper is the conceptualization of project tailoring as a knowledge legitimacy process and the development of a process model that explains how project decisions emerge in hybrid intelligence environments. The paper contributes to project studies, knowledge management, and organizational research by explaining project decision making as a culturally and technologically embedded knowledge process.

7 Acknowledgements

The research is supported by the Research Centre at the Faculty of Business and Economics (No. PE-GTK-GSKK A095000000-13), University of Pannonia (Veszprém, Hungary).

8 Ethics Declaration

This paper is conceptual in nature and does not involve human participants, personal data, or any form of empirical data collection. Therefore, ethical approval was not required for this study.

9 AI Declaration

AI-based tool (ChatGPT) was used to support literature search, language editing, and manuscript structuring. No AI tools were used to develop the conceptual model or the theoretical arguments. The author takes full responsibility for the content of this paper.

References

Abimbola, S., van de Kamp, J., Lariat, J., Rathod, L., Klipstein-Grobusch, K., van der Graaf, R. and Bhakuni, H. (2024) 'Unfair knowledge practices in global health: A realist synthesis', *Health Policy and Planning*, Vol. 39, No. 6, pp. 636-650. https://doi.org/10.1093/heapol/czae030

Adeinat, I. and Abdulfatah, F. (2019) 'Organizational culture and knowledge management processes: Case study in a public university', *VINE Journal of Information and Knowledge Management Systems*, Vol. 49, No. 1, pp. 35-53. https://doi.org/10.1108/VJIKMS-05-2018-0041

Antonova, A, (2025) 'The Future of Work and the Jobs in Transition: Modelling Scenarios and Implications for Organisational Culture', in Csepregi, A., Benedek, P. (eds.) *The Multifaceted World of Organizational Culture: Research Directions and Corporate Experiences.* Book of Conference Proceedings, Wekerle International University, Budapest, pp. 4-14.

Andrews, L.M., Munaretto, S., Mees, H.L.P. and Driessen, P.P.J. (2024) 'Conceptualising boundary work activities to enhance credible, salient and legitimate knowledge', *Environmental Science & Policy*, Vol. 155, Article 103722. https://doi.org/10.1016/j.envsci.2024.103722

Araki, M. and Martins, H. (2023) 'Creativity governance: A conceptual framework for tailoring governance to the creativity and uncertainty in entrepreneurial projects', *European Management Review*, Vol. 21, No. 3, pp. 701-713. https://doi.org/10.1111/emre.12607

Brill, D., Schnugg, C. and Stary, C. (2025) 'Collaborative construction of meaning: facilitating sensemaking moments through aesthetic knowledge generation', *Knowledge Management Research & Practice*, Vol. 23, No. 4, pp. 382–396. https://doi.org/10.1080/14778238.2024.2382828

Cicmil, S., Williams, T., Thomas, J. and Hodgson, D. (2006) 'Rethinking project management: researching the actuality of projects', *International Journal of Project Management*, Vol. 24, No. 8, pp. 675–686. https://doi.org/10.1016/j.ijproman.2006.08.006

Clegg, S., Killen, C.P., Biesenthal, C. and Sankaran, S. (2018) 'Practices, projects and portfolios: Current research trends and new directions', *International Journal of Project Management*, Vol. 36, No. 5, pp. 762-772. https://doi.org/10.1016/j.ijproman.2018.03.008

de Metz, N., Jansen van Rensburg, M. and Davis, A. (2024) 'Shared sensemaking during a strategic change process: a non-managerial perspective', *Journal of Strategy and Management*, Vol. 17, No. 4, pp. 763–780. https://doi.org/10.1108/JSMA-10-2023-0264

Deephouse, D. L., Bundy, J., Tost, L. P. and Suchman, M. C (2017) 'Organizational legitimacy: Six key questions', in Greenwood, R., Oliver, C., Lawrence, T. and Meyer, R. (eds.) *The SAGE Handbook of Organizational Institutionalism*. Vol. 4, No. 2, pp. 27-54

Fabri, L., Häckel, B., Oberländer, A., Rieg, M. and Stohr, A. (2023) 'Disentangling human–AI hybrids: Conceptualizing the interworking of humans and AI-enabled systems', *Business & Information Systems Engineering*, Vol. 65, No. 6, pp. 623–641. https://doi.org/10.1007/s12599-023-00810-1

Harris, S.G. (1994) 'Organizational culture and individual sensemaking: A schema-based perspective', *Organization Science*, Vol. 5, No. 3, pp. 309–321. https://doi.org/10.1287/orsc.5.3.309

Jaakkola, E. (2020) 'Designing conceptual articles: four approaches', *AMS review*, Vol. 10, No. 1, pp. 18-26. https://doi.org/10.1007/s13162-020-00161-0

Jansson, N., Lunkka, N., Suhonen, M., Meriläinen, M. and Wiik, H. (2020) 'The emergence of sensemaking through sociomaterial practices', *Journal of Organizational Change Management*, Vol. 33, No. 4, pp. 597–607. https://doi.org/10.1108/JOCM-10-2018-0280

Jarrahi, M.H. (2018) 'Artificial intelligence and the future of work: Human–AI symbiosis in organizational decision making', *Business Horizons*, Vol. 61, No. 4, pp. 577–586. https://doi.org/10.1016/j.bushor.2018.03.007

Ledro, C., Nosella, A. and Dalla Pozza, I. (2023) 'Integration of AI in CRM: Challenges and guidelines', *Journal of Open Innovation: Technology, Market, and Complexity*, Vol. 9, No. 4, Article 100151. https://doi.org/10.1016/j.joitmc.2023.100151

Long, D.W. and Fahey, L. (2000) 'Diagnosing cultural barriers to knowledge management', *Academy of Management Perspectives*, Vol. 14, No. 4, pp. 113–127. https://doi.org/10.5465/ame.2000.3979820

Lundin, R.A. and Söderholm, A. (1995) 'A theory of the temporary organization', *Scandinavian Journal of Management*, Vol. 11, No. 4, pp. 437–455. https://doi.org/10.1016/0956-5221(95)00036-U

MacInnis, D. J. (2011) 'A framework for conceptual contributions in marketing', *Journal of Marketing*, Vol. 75, No. 4, pp. 136-154. https://doi.org/10.1509/jmkg.75.4.136

Maitlis, S. and Christianson, M. (2014) 'Sensemaking in organizations: Taking stock and moving forward', *Academy of Management Annals*, Vol. 8, No. 1, pp. 57–125. https://doi.org/10.5465/19416520.2014.873177

Miterev, M., Mancini, M. and Turner, J.R. (2017a) 'Towards a design for the project-based organization', *International Journal of Project Management*, Vol. 35, No. 3, pp. 479–491. https://doi.org/10.1016/j.ijproman.2016.12.007

Olan, F. Arakpogun, E.O., Suklan, J., Nakpodia, F., Damij, N. and Jayawickrama, U. (2022) 'Artificial intelligence and knowledge sharing: Contributing factors to organizational performance', *Journal of Business Research*, Vol. 145, pp. 605–615. https://doi.org/10.1016/j.jbusres.2022.03.008

Orlikowski, W.J. (2007) 'Sociomaterial practices: Exploring technology at work', *Organization Studies*, Vol. 28, No. 9, pp. 1435–1448. https://doi.org/10.1177/0170840607081138

Packendorff, J. (1995) 'Inquiring into the temporary organization: New directions for project management research', *Scandinavian Journal of Management*, Vol. 11, No. 4, pp. 319–333. https://doi.org/10.1016/0956-5221(95)00018-Q

Pagnozzi, F., Pescatore, I., Cartone, A. and Antonelli, G. (2024) 'Organisational Culture and Hybrid Work in the Italian Public Sector—The Case of a Small University, in Agrifoglio, R., Lazazzara, A. and Za, S (eds.) *Navigating Digital Transformation: Organizational Change, Digital Work, and Individual Behavior*, Springer, Cham, pp. 177–194. https://doi.org/10.1007/978-3-031-76970-2_12

Pescatore, I., Pagnozzi, F. and Antonelli, G. (2022) 'Analysing the bottom-up approach to develop organisational culture in virtualised organisations', in Lazazzara, A., Reina, R., Za, S. (eds) *Towards Digital and Sustainable Organisations. ItAIS 2022. Lecture Notes in Information Systems and Organisation*, Springer Nature Switzerland, Cham, pp. 141–154. https://doi.org/10.1007/978-3-031-52880-4_9

Sandberg, J. and Tsoukas, H. (2015) 'Making sense of the sensemaking perspective: Its constituents, limitations, and opportunities for further development', *Journal of Organizational Behavior*, Vol. 36(S1), pp. S6–S32. https://doi.org/10.1002/job.1937

Schein, E.H. (2010) *Organizational culture and leadership*, 4th ed., Wiley, Hoboken, NJ.

Sherson, J. (2024) 'A hybrid intelligent change management approach to generative AI adoption', in *HHAI 2024: Hybrid Human–AI Systems for the Social Good*, IOS Press, Amsterdam, pp. 472–474. https://doi.org/10.3233/FAIA240228

Shrestha, Y.R., Ben-Menahem, S.M. and von Krogh, G. (2019) 'Organizational decision-making structures in the age of artificial intelligence', *California Management Review*, Vol. 61, No. 4, pp. 66–83. https://doi.org/10.1177/0008125619862257

Sloot, R.N.F., Troje, D., Voordijk, H. and Volker, L. (2024) 'Change in a project-based organization: The mutual shaping of institutional logics and change programs', *International Journal of Project Management*, Vol. 42, No. 3, Article 102589. https://doi.org/10.1016/j.ijproman.2024.102589

Söderlund, J. (2011) 'Pluralism in project management: Navigating the crossroads of specialization and fragmentation', *International Journal of Management Reviews*, Vol. 13, No. 2, pp. 153–176. https://doi.org/10.1111/j.1468-2370.2010.00290.x

Sydow, J. and Windeler, A. (2020) 'Temporary organizing and permanent contexts', *Current Sociology*, Vol. 68, No. 4, pp. 480–498. https://doi.org/10.1177/0011392120907629

Wei, Y. and Miraglia, S. (2017) 'Organizational culture and knowledge transfer in project-based organizations: Theoretical insights from a Chinese construction firm', *International Journal of Project Management*, Vol. 35, No. 4, pp. 571–585. https://doi.org/10.1016/j.ijproman.2017.02.010

Whetten, D. A. (1989) 'What constitutes a theoretical contribution?' *Academy of management review*, Vol. 14, No. 4, pp. 490-495. https://doi.org/10.5465/amr.1989.4308371

Whittle, A., Vaara, E. and Maitlis, S. (2023) 'The role of language in organizational sensemaking: An integrative theoretical framework and an agenda for future research', *Journal of Management*, Vol. 49, No. 6, pp. 1807-1840. https://doi.org/10.1177/01492063221147295

Rethinking Help-Seeking in AI-Mediated Workplaces: A Socio-Technical and Resource-Based Perspective

Filomena Pagnozzi[1], Zhijie (Helen) Zhao[2]
[1]University of G. d'Annunzio, Chieti-Pescara, Italy
[2]Amsterdam Business School, University of Amsterdam, Amsterdam, The Netherlands

Abstract

Artificial intelligence (AI) systems are increasingly embedded in contemporary workplaces, offering employees new sources of assistance for obtaining information and receiving support during work tasks. This transformation is reshaping how help is accessed, as employees can now turn not only to colleagues and supervisors but also to AI-based tools such as chatbots and generative AI systems.

However, the help-seeking literature has traditionally conceptualized this behavior as inherently interpersonal, shaped by relational dynamics such as reputational concerns and power relations. While this perspective has generated important insights, it offers only a partial account of help-seeking in AI-mediated workplaces, where sources of assistance extend beyond human interactions.

In these contexts, help-seeking decisions are shaped by the interaction between socio-technical infrastructures, including availability, accessibility, characteristics of AI systems and individuals' evaluations of potential resource gains and losses related to help sources. Despite the relevance of this shift, existing research lacks a framework that conceptualizes help-seeking as a multi-level phenomenon emerging from the interaction of organizational and individual factors.

To address this gap, this paper integrates Socio-Technical Systems theory and Conservation of Resources theory to examine how organizational and individual antecedents jointly shape help-seeking behaviors. Drawing on this integrated perspective, we propose a conceptual framework that identifies organizational infrastructure conditions and individual intrapersonal and interpersonal factors of three help-seeking strategies: AI-based, human-based, and hybrid help-seeking.

By combining these perspectives, we reconceptualize help-seeking in AI-mediated workplaces as a multi-level resource-regulation process and provide a foundation for future research on help-seeking in AI-mediated workplaces.

Keywords: AI-mediated work; Help-seeking behavior; Socio-Technical Systems; Conservation of Resources

Paper type: Academic Research Paper

1 Introduction

In contemporary workplaces, the way employees seek and obtain help is quietly but changing. For a long time, when individuals encountered difficulties in their work tasks, they turned to humans (i.e. supervisors, colleagues) to access the knowledge and support needed to move forward. Within this traditional view, help-seeking has been understood as an interpersonal process—one that unfolds through social relationships and is shaped by concerns about reputation, power dynamics, and norms of reciprocity (Lee, 1997; Bamberger, 2009; Hargadon and Bechky, 2006; Pagnozzi et al., 2024). Asking for help was therefore not a neutral act: it often meant exposing one's limitations, navigating social expectations, and weighing the benefits of receiving support against the costs of appearing incompetent or dependent in front of others (Nadler et al., 2003; Lee, 1997).

This understanding is being challenged by the presence of Artificial Intelligence (AI) in organizational settings. Increasingly, employees can rely not only on other people but also on AI-based systems—such as chatbots and generative AI tools—to obtain assistance. These technologies provide support that is immediate, scalable, and often anonymous, making it possible to seek help without the social exposure typically associated with interpersonal interactions. In many cases, AI tools are already being used as everyday problem-solving companions or personal assistants, offering on-demand guidance that is independent of others' availability or willingness to help (Meyer von Wolff et al., 2020; Gkinko and Elbanna, 2023; Brynjolfsson et al., 2023; Noy and Zhang, 2023).

As a consequence, help-seeking can no longer be seen as purely a social act. It is becoming a hybrid behavior that takes shape at the intersection of human and technological interactions. This shift raises new and important questions: why do individuals choose to seek help from AI rather than from other people? How do they evaluate the costs and benefits associated with these different sources of support? And how do organizational contexts and technological infrastructures influence these choices?

Despite the relevance of these questions, existing research has yet to capture this transformation. Studies on help-seeking continue to emphasize interpersonal dynamics (Bamberger, 2009), often overlooking the role

of technological systems. At the same time, research on AI adoption tends to focus on usage and performance outcomes, paying less attention to the behavioral and motivational processes that underlie decisions to seek help (Davis, 1989; Venkatesh et al., 2003). More broadly, although socio-technical scholars have long argued for the need to jointly consider social and technological dimensions (Trist and Bamforth, 1951; Bostrom and Heinen, 1977; Pasmore et al., 2019), this integration remains largely absent in the study of help-seeking.

Against this backdrop, we propose that help-seeking in AI-mediated workplaces can be understood as a multi-level resource regulation process, shaped by the interplay between socio-technical conditions and individual resource evaluations. To do so, we develop an integrated framework that combines Socio-Technical Systems (STS) theory (Bednar and Welch, 2020; Pagnozzi and Antonelli, 2025) with Conservation of Resources (COR) theory (Hobfoll, 1989; 2001). This approach allows us to capture both the structural context in which help-seeking takes place and the underlying motivations that drive individuals' decisions. By bringing these perspectives together, the study contributes to a more nuanced understanding of help-seeking in contemporary organizations, where the boundaries between human and technological sources of support are becoming increasingly blurred.

2 Theoretical Background

2.1 Help-Seeking as an Interpersonal Behavior

Help-seeking is a fundamental process through which individuals address work-related challenges and acquire knowledge to perform tasks effectively. It is defined as requesting assistance from others to solve problems (Lee, 1997; Bamberger, 2009). In organizations, it is inherently relational, involving interactions between help-seekers and help-givers within social and organizational networks (Hargadon and Bechky, 2006; Mueller and Kamdar, 2011). Research has examined help-seeking mainly through social and psychological perspectives, showing that individuals consider not only task needs but also social implications. Asking for help may signal lack of competence or autonomy, threatening image, status, or power (Nadler et al., 2003; Lee, 1997). Thus, help-seeking reflects a trade-off between expected benefits and perceived interpersonal costs (Bamberger, 2009). Help-seeking is also linked to positive outcomes, including learning, problem-solving, and performance (Mueller and Kamdar, 2011), as well as stronger relationships, trust, and collaboration (Hargadon and Bechky, 2006). However, it remains a complex and potentially risky behavior shaped by individual and organizational factors. Despite these insights, prior research largely treats help-seeking as socially embedded, focusing on interpersonal dynamics. This view is partial in contemporary organizations. First, it assumes support is provided by individuals, overlooking technological systems. Yet digital infrastructures—including platforms, knowledge management systems, and AI-based tools—shape how employees access expertise (Leonardi and Barley, 2010; Faraj and Azad, 2012). Second, organizational conditions are underexplored. Leadership and culture can enable or constrain help-seeking: supportive leadership fosters psychological safety (Edmondson, 1999; Detert and Burris, 2007; Carmeli et al., 2010), while collaborative cultures promote knowledge sharing, unlike competitive environments (Schein, 2010; Van Dyne et al., 2003; Bamberger, 2009). Finally, motivational processes remain underexplored. Although individuals evaluate costs and benefits, it is unclear how these evaluations unfold when both human and technological support are available (Nadler et al., 2003; Bamberger, 2009). These limitations call for an integrated perspective on individual, organizational, and technological factors.

2.2 The Emergence of AI-Mediated Help

The growing diffusion of AI in organizations highlights the limitations of traditional views of help-seeking by introducing new forms of support. AI-based systems—such as chatbots, recommendation engines, and generative tools—are increasingly embedded in everyday work, offering alternative sources of assistance that differ from human interactions. Unlike traditional help providers, AI systems offer immediate, continuously available, and scalable support. This reduces constraints associated with interpersonal help-seeking, including availability, time coordination, and social risk. Studies show that generative AI tools enhance productivity and support complex problem-solving, improving access to assistance (Brynjolfsson et al., 2023; Noy and Zhang, 2023). Interacting with AI in less socially exposed contexts may also lower psychological barriers to asking for help (Gkinko and Elbanna, 2023). At the same time, AI-mediated help-seeking introduces new forms of uncertainty. Individuals must evaluate the reliability, usefulness, and quality of AI outputs, as well as their ability to interact effectively with these systems. Research on technology adoption highlights the role of trust, perceived usefulness, and ease of use in shaping behavior (Davis, 1989; Venkatesh et al., 2003; Glikson and

Woolley, 2020). Thus, turning to AI is not simply a substitute for human help but involves distinct cognitive and behavioral considerations. More broadly, the integration of AI reflects a transformation of work systems. Tools such as chatbots are reshaping help-seeking practices by requiring new skills, redefining roles, and influencing organizational structures and cultures.

3 Theoretical Foundation

3.1 Socio-Technical Systems (STS) Theory

STS theory provides a valuable lens for understanding help-seeking in organizations shaped by technological mediation. Originating from Trist and Bamforth (1951), STS posits that organizational outcomes emerge from the interaction between two interdependent subsystems: the social subsystem (individuals, relationships, structures) and the technical subsystem (tools, technologies, processes, tasks) (Bostrom and Heinen, 1977; Cherns, 1987; Mumford, 2006). Rather than treating these as separate, STS emphasizes their mutual interdependence: technological changes reshape social dynamics, while human needs and organizational arrangements influence how technologies are designed and used (Pasmore et al., 2019; Bednar and Welch, 2020; Leonardi, 2011). Applied to help-seeking, this perspective suggests that the behavior is not purely interpersonal but embedded in a broader socio-technical system. While prior research has focused on social dimensions such as trust and reciprocity, technological infrastructures also shape when, how, and from whom help is sought (Faraj and Azad, 2012; Leonardi and Barley, 2010; Pagnozzi and Antonelli, 2025). Digital platforms, knowledge management systems, and AI-based tools influence the visibility of expertise, access to support, and the timing and modality of help requests. AI systems, such as chatbots and generative tools, represent a key evolution of the technical subsystem. They provide new sources of assistance while reshaping interaction patterns, expectations, and organizational norms (Brynjolfsson and McAfee, 2014; Raisch and Krakowski, 2021). These tools offer immediate and scalable support but also require new skills and forms of coordination (Noy and Zhang, 2023; Pagnozzi et al., 2024).From an STS perspective, this transformation is socio-technical, affecting culture, work design, and how individuals engage with human and non-human sources of help (Orlikowski, 2007; Leonardi, 2011). Help-seeking in AI-mediated workplaces thus emerges from the interplay between social structures and technological affordances. However, while STS explains the contextual conditions of help-seeking, it does not fully account for individual motivations and decision-making processes.

3.2 Conservation of Resources (COR) Theory

COR theory provides a complementary lens to explain the individual-level motivations underlying help-seeking in AI-mediated workplaces. Developed by Hobfoll (1989), it posits that individuals seek to acquire, retain, and protect resources—broadly defined as anything valued for achieving goals. Central to the theory is the tension between resource acquisition and conservation: individuals evaluate whether investing resources such as time, effort, and social capital will generate sufficient gains to justify the cost (Hobfoll et al., 2018; Halbesleben et al., 2014). From this perspective, help-seeking is a resource investment strategy, where individuals expend resources to obtain greater returns (Lim et al., 2020). Engagement in help-seeking depends not only on need, but on the evaluation of expected gains relative to potential losses, shaped by individuals' current resource state and prior experiences (Hobfoll et al., 2018). In AI-mediated workplaces, this evaluation involves new trade-offs. The primacy of loss principle suggests that individuals are particularly sensitive to the costs of help-seeking, including effort, social exposure, and psychological risks such as appearing incompetent or incurring indebtedness (Halbesleben et al., 2014). AI-based assistance alters this cost structure by reducing social visibility, thereby lowering psychological barriers and enabling help-seeking that might otherwise be avoided (Bamberger, 2009; Lim et al., 2020). At the same time, the resource investment principle suggests that individuals with higher AI self-efficacy or prior positive experiences are more likely to engage in AI-based help-seeking, as they expect greater returns from their investment (Hobfoll et al., 2018; Lim et al., 2020).

4 Integrating STS and COR: Toward a Multi-Level Perspective

4.1 Limitations of Single-Theory Explanations

STS and COR theories each offer valuable but partial insights into help-seeking in AI-mediated workplaces. STS explains how social and technical subsystems shape the conditions under which help-seeking becomes possible or constrained (Trist and Bamforth, 1951; Pasmore et al., 2019; Bednar and Welch, 2020), but it does not fully account for individual decision-making. COR theory, by contrast, explains individual motivations and

differences in help-seeking based on resource dynamics (Hobfoll, 1989; Halbesleben et al., 2014), yet overlooks the technological and organizational context in which these evaluations occur. Taken separately, neither perspective fully explains help-seeking in AI-mediated workplaces.

4.2 Why an Integrated Perspective is Needed

Help-seeking in AI-mediated workplaces is inherently multi-level and cannot be explained by individual or contextual factors alone. Integrating STS and COR enables a more complete understanding of three key issues. First, why individuals choose AI or human help: this depends on the interaction between available socio-technical conditions and individual resource evaluations (Hobfoll et al., 2018; Pasmore et al., 2019). Second, why the same AI system produces different behaviors: STS highlights variation in organizational embedding, while COR explains differences in individual resource states. Third, how organizational contexts shape decisions, as socio-technical conditions define the environment within which individuals evaluate costs and benefits (Orlikowski, 2007; Leonardi, 2011). This integration provides a more comprehensive account of help-seeking in AI-mediated workplaces.

4.3 Help-Seeking as a Multi-level Resource Regulation Process

Building on this integration, we conceptualize help-seeking in AI-mediated workplaces as a multi-level resource regulation process arising from the interaction between socio-technical infrastructures and individual resource motives. Organizational infrastructures—including the design, availability, and embedding of AI tools—define what forms of assistance are accessible, at what cost, and under which norms. Within this environment, individuals evaluate whether seeking help from AI or human sources is likely to generate resource gains or losses, considering effort, social exposure, and potential image risks (Hobfoll et al., 2018; Halbesleben et al., 2014; Lim et al., 2020). Help-seeking thus emerges from the ongoing interplay between organizational conditions and individual acquisition–conservation orientations.

5 Conceptual Framework of Help-Seeking in AI-Mediated Workplaces

Building on the integrated STS-COR perspective, we propose that help-seeking in AI-mediated workplaces is shaped by a set of organizational and individual factors that influence how individuals evaluate the relative costs and benefits of seeking help from AI versus human sources. From this perspective, help-seeking can be conceptualized as a choice among alternative resource-seeking strategies, in which individuals select the source of assistance that maximizes expected resource gains while minimizing resource losses.

5.1 Help-Seeking Behaviors

We conceptualize help-seeking as a set of alternative resource-seeking strategies through which individuals address work-related challenges (Bamberger, 2009; Nadler et al., 2003). These strategies differ in their cost–benefit profiles and reflect different ways of navigating the socio-technical environment.

AI-based help-seeking occurs when individuals rely primarily on AI systems. This strategy involves lower social costs and higher accessibility but may entail uncertainty about output quality and limited contextual understanding in complex situations (Glikson and Woolley, 2020; Raisch and Krakowski, 2021).

Human-based help-seeking involves seeking assistance from colleagues or supervisors. It provides richer, context-sensitive support and relational benefits but entails higher social and psychological costs, such as impression management and indebtedness (Nadler et al., 2003; Bamberger, 2009).

Hybrid help-seeking combines AI and human sources, allowing individuals to balance their advantages and limitations. This aligns with perspectives on human–AI collaboration emphasizing their complementarity (Jarrahi, 2018).

In line with the STS-COR perspective, these strategies emerge from the interplay between organizational infrastructures and individual motivations. Individuals tend to select the option that maximizes expected gains while minimizing perceived costs.

5.2 Organizational Infrastructures

Organizational infrastructures define the socio-technical environment within which help-seeking decisions are made. These infrastructures shape the structural conditions of action by influencing the accessibility, perceived legitimacy, and expected value of different sources of assistance (Faraj and Azard, 2012; Leonardi, 2011).

First, AI accessibility refers to the availability and integration of AI tools in everyday work. High levels of accessibility reduce the effort and time required to seek help, thereby lowering the resource costs associated with AI-based help-seeking (Gkinko and Elbanna, 2023).

Second, technological affordances capture the capabilities of AI systems to provide accurate, timely, and task-relevant support (Leonardi, 2011; Orlikowski, 2007). Strong affordances increase the perceived usefulness of AI, enhancing the expected resource gains from engaging with these systems (Davis, 1989; Vankatesh et al., 2013; Pagnozzi et al., 2024).

Third, organizational norms shape the social acceptability of help-seeking. In collaborative environments, both human and AI-based help are supported, while in competitive contexts individuals may prefer AI as it reduces social exposure (Schein, 2010; Van Dyne et al., 2003; Bamberger, 2009).

Finally, AI legitimacy reflects the extent to which its use is considered appropriate and valued. When legitimized, the reputational and psychological costs decrease, making AI-based or hybrid help-seeking more attractive (Martin and Waldman, 2023; Scott, 2013).

5.3 Individual Motivations

While organizational infrastructures define the external conditions of action, individual factors shape how these are interpreted. In line with COR theory, individuals differ in their sensitivity to resource gains and losses and in their willingness to seek help from different sources (Hobfoll, 1989; Hobfoll et al., 2018). Importantly, these factors influence not only whether individuals seek help, but also which source they choose.

Intrapersonal factors shape preferences toward help-seeking. Individuals with a high learning goal orientation are more likely to prioritize resource acquisition and engage in help-seeking, including through AI (Lim et al., 2020). Openness to experience increases experimentation with new technologies, while trust in AI enhances expectations about its usefulness and reliability (McGrath et al., 2025; Wanberg and Kammeyer-Mueller, 2000).

Interpersonal factors influence sensitivity to social costs. High performance goal orientation increases concerns about appearing incompetent, making individuals less likely to seek help from others and more likely to rely on AI (VandeWalle, 1997). Similarly, avoidance of indebtedness and relational conflicts increase the perceived costs of asking others for help, leading individuals to prefer AI-based support (Chan, 2013).

In line with the STS-COR perspective, the three help-seeking strategies emerge from the interplay between organizational infrastructures and individual motivations. Organizational infrastructures define available sources, while individual factors shape their evaluation, leading to preferences for AI-based, human-based, or hybrid help-seeking. This framework is summarized in Figure 1.

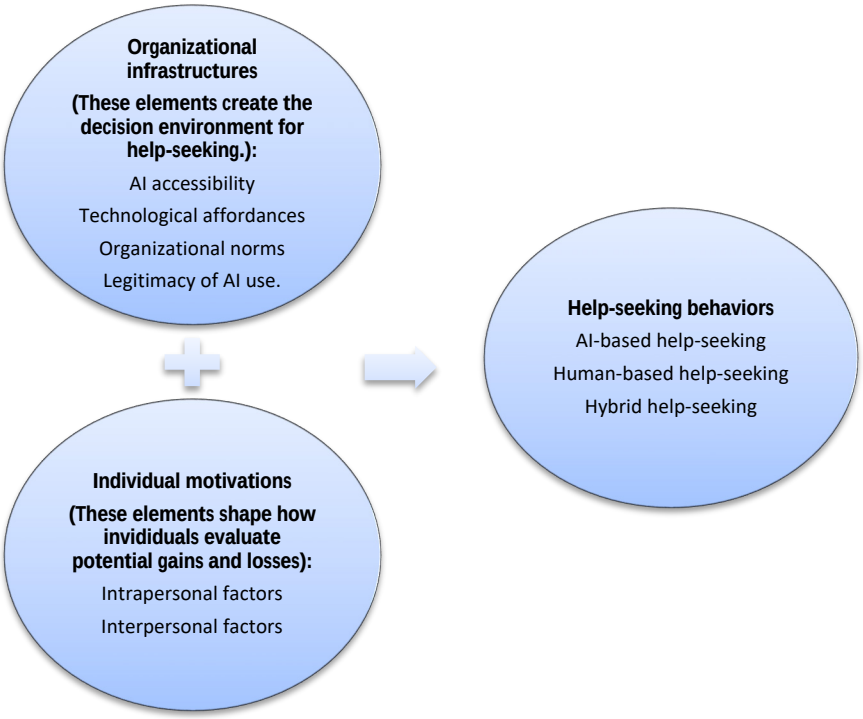


Figure 1 Conceptual Framework of Help-Seeking in AI-Mediated Workplaces

6 Discussion

6.1 Theoretical Implications

This paper offers theoretical contributions to the help-seeking literature. First, we extend the help-seeking literature by challenging its dominant view as an inherently interpersonal behavior (Bamberger, 2009), showing that, in AI-mediated workplaces, help-seeking is better understood as a socio-technical and multi-source decision process. While STS theory explains how socio-technical conditions shape the context of action (Trist and Bamforth, 1951; Pasmore et al., 2019), and COR theory explains how individuals evaluate resource gains and losses (Hobfoll, 1989; Hobfoll et al., 2018), we show that help-seeking emerges from their joint interaction. In doing so, we provide a more complete account of help-seeking that captures both structural conditions and individual motivations (Halbesleben et al., 2014; Lim et al., 2020).

While prior research has largely focused on whether individuals seek help (Nadler et al., 2003; Bamberger, 2009), our framework explains how employees choose between AI-based, human-based, and hybrid help-seeking when multiple sources are available (Chen et al., 2024; Lim et al., 2020). By organizing antecedents into organizational infrastructures and individual motivations, we identify the conditions under which different help-seeking strategies are more likely to emerge. This contributes to a more nuanced understanding of help-seeking in contemporary organizations where employees navigate between human and technological sources of support (Chen et al., 2024; Lim et al., 2020).

6.2 Practical Implications

This paper also offers practical implications for managing work in AI-integrated environments. Organizations should not rely on technological deployment alone, but actively design the socio-technical conditions surrounding AI tools. Factors such as accessibility, organizational norms, and the legitimacy of AI use shape how employees perceive and engage with different help sources (Leonardi, 2011; Faraj and Azad, 2012). Managers and HR practitioners should should pay attention to reduce the perceived costs and increase the expected benefits of using AI where appropriate.

At the same time, employees do not respond uniformly to the availability of AI tools. Individual factors such as goal orientation and trust in AI shape how employees evaluate and engage with AI as a help source. This heterogeneity implies that organizations should invest in targeted support, such as AI literacy programs and psychologically safe environments, to help employees develop the confidence needed to engage effectively with AI tools.

Finally, AI and human help should not be treated as substitutes. but as complementary resources. Encouraging norms that support both forms of help-seeking can enable hybrid strategies that leverage the strengths of human and AI collaboration (Jarrahi, 2018).

6.3 Future Directions and Conclusion

This paper is conceptual in nature and the framework we propose calls for empirical validation across multiple methodological approaches. Quantitative studies could test the relationships between individual antecedents and help-seeking source preferences through survey-based or experimental designs, examining for instance how organizational and individual factors predict AI-based, human-based, or hybrid help-seeking. Qualitative approaches, including interviews and ethnographic observation, would complement this by capturing how individuals navigate and make sense of help-seeking decisions in AI-mediated workplaces in practice, as well as how organizational infrastructures shape these processes over time.

Beyond methodology, our conceptual framework invites further refinement. Adopting a dynamic perspective could help explain how help-seeking patterns evolve as employees accumulate experience with AI tools, and whether early interaction experiences create trajectories of increasing AI reliance or continued preference for human support. Furthermore, future work could examine the outcomes of different help-seeking strategies, exploring whether AI-based, human-based, and hybrid help-seeking differ in their effects on individual performance and well-being, as well as on collective knowledge sharing and team collaboration.

In conclusion, the way employees seek and obtain help—once understood as an inherently interpersonal process—is now being fundamentally redefined. Help-seeking can no longer be understood as a purely social act, but a socio-technical and multi-level process, in which individual motivations and socio-technical conditions jointly determine how employees choose between different sources of support. By integrating STS and COR theories, we reconceptualize help-seeking as a resource regulation process in which individuals evaluate the costs and benefits of alternative help sources within a socio-technical environment. In doing so, we shift the

focus from whether individuals seek help to how they select among AI-based, human-based, and hybrid forms of assistance. This offers a foundation for future research on help-seeking in AI-mediated workplaces.

7 Ethics Declaration

This paper is a conceptual study and does not involve the collection or analysis of empirical data from human participants. Therefore, no ethical approval was required.

8 AI Declaration

Claude (Anthropic), an AI language model, was used for linguistic refinement of the manuscript. All AI-assisted content was critically reviewed by the authors. The use of AI was limited to language editing and did not extend to the development of ideas or the conceptual framework, which remain the authors' original work.

References

Bamberger, P. (2009) 'Employee help-seeking: Antecedents, consequences and new insights for future research', *Research in Personnel and Human Resources Management*, Vol. 28, pp. 49–98.

Bednar, P.M. and Welch, C. (2020) 'Socio-technical perspectives on smart working: Creating meaningful and sustainable systems', *Information Systems Frontiers*, Vol. 22, No. 2, pp. 281–298.

Bostrom, R.P. and Heinen, J.S. (1977) 'MIS problems and failures: A socio-technical perspective: Part II: The application of socio-technical theory', *MIS Quarterly*, Vol. 1, No. 4, pp. 11–28.

Brynjolfsson, E. and McAfee, A. (2014) *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*, New York: W.W. Norton and Company.

Brynjolfsson, E., Li, D. and Raymond, L.R. (2023) 'Generative AI at work', *NBER Working Paper*, No. 31161.

Carmeli, A., Reiter-Palmon, R. and Ziv, E. (2010) 'Inclusive leadership and employee involvement in creative tasks in the workplace: The mediating role of psychological safety', *Creativity Research Journal*, Vol. 22, No. 3, pp. 250–260.

Chan, M.E. (2013) 'Antecedents of instrumental interpersonal help-seeking: An integrative review', *Applied Psychology*, Vol. 62, No. 4, pp. 571–596.

Chen, K., Chen, G., Wu, Q., Liu, W. and Zhao, H. (2024) 'Help-seeking at work: An integrative review, organizing framework and agenda for future research', *Nankai Business Review International*, Vol. 15, No. 2, pp. 351–374.

Cherns, A. (1987) 'Principles of sociotechnical design revisited', *Human Relations*, Vol. 40, No. 3, pp. 153–162.

Davis, F.D. (1989) 'Perceived usefulness, perceived ease of use, and user acceptance of information technology', *MIS Quarterly*, Vol. 13, No. 3, pp. 319–340.

Detert, J.R. and Burris, E.R. (2007) 'Leadership behavior and employee voice: Is the door really open?', *Academy of Management Journal*, Vol. 50, No. 4, pp. 869–884.

Edmondson, A. (1999) 'Psychological safety and learning behavior in work teams', *Administrative Science Quarterly*, Vol. 44, No. 2, pp. 350–383.

Faraj, S. and Azad, B. (2012) 'The materiality of technology: An affordance perspective', in Leonardi, P.M., Nardi, B.A. and Kallinikos, J. (eds.) *Materiality and Organizing: Social Interaction in a Technological World*, Oxford: Oxford University Press, pp. 237–258.

Gkinko, L. and Elbanna, A. (2023) 'The appropriation of conversational AI in the workplace: A taxonomy of AI chatbot users', *International Journal of Information Management*, Vol. 69, p. 102568.

Glikson, E. and Woolley, A.W. (2020) 'Human trust in artificial intelligence: Review of empirical research', *Academy of Management Annals*, Vol. 14, No. 2, pp. 627–660.

Halbesleben, J.R.B., Neveu, J.P., Paustian-Underdahl, S.C. and Westman, M. (2014) 'Getting to the "COR": Understanding the role of resources in conservation of resources theory', *Journal of Management*, Vol. 40, No. 5, pp. 1334–1364.

Hargadon, A.B. and Bechky, B.A. (2006) 'When collections of creatives become creative collectives: A field study of problem solving at work', *Organization Science*, Vol. 17, No. 4, pp. 484–500.

Hobfoll, S.E. (1989) 'Conservation of resources: A new attempt at conceptualizing stress', *American Psychologist*, Vol. 44, No. 3, pp. 513–524.

Hobfoll, S.E. (2001) 'The influence of culture, community, and the nested-self in the stress process: Advancing conservation of resources theory', *Applied Psychology*, Vol. 50, No. 3, pp. 337–421.

Hobfoll, S.E., Halbesleben, J., Neveu, J.P. and Westman, M. (2018) 'Conservation of resources in the organizational context: The reality of resources and their consequences', *Annual Review of Organizational Psychology and Organizational Behavior*, Vol. 5, pp. 103–128.

Jarrahi, M.H. (2018) 'Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making', *Business Horizons*, Vol. 61, No. 4, pp. 577–586.

Lee, F. (1997) 'When the going gets tough, do the tough ask for help? Help-seeking and power motivation in organizations', *Organizational Behavior and Human Decision Processes*, Vol. 72, No. 3, pp. 336–363.

Leonardi, P.M. (2011) 'When flexible routines meet flexible technologies: Affordance, constraint, and the imbrication of human and material agencies', *MIS Quarterly*, Vol. 35, No. 1, pp. 147–167.

Leonardi, P.M. and Barley, S.R. (2010) 'What's under construction here? Social action, materiality, and power in constructivist studies of technology and organizing', *Academy of Management Annals*, Vol. 4, No. 1, pp. 1–51.

Lim, J.H., Tai, K., Bamberger, P.A. and Morrison, E.W. (2020) 'Soliciting resources from others: An integrative review', *Academy of Management Annals*, Vol. 14, No. 1, pp. 122–159.

Martin, K. and Waldman, A. (2023) 'Are algorithmic decisions legitimate? The effect of process and outcomes on perceptions of legitimacy of AI decisions', *Journal of Business Ethics*, Vol. 183, No. 3, pp. 653–670.

McGrath, M.J., Lack, O., Tisch, J. and Duenser, A. (2025) 'Measuring trust in artificial intelligence: Validation of an established scale and its short form', *Frontiers in Artificial Intelligence*, Vol. 8, p. 1582880.

Meyer von Wolff, R., Hobert, S., Masuch, K. and Schumann, M. (2020) 'Chatbots at digital workplaces: A grounded-theory approach for surveying application areas and objectives', *Pacific Asia Journal of the Association for Information Systems*, Vol. 12, No. 2, p. 3.

Mueller, J.S. and Kamdar, D. (2011) 'Why seeking help from teammates is a blessing and a curse: A theory of help-seeking and individual creativity in team contexts', *Journal of Applied Psychology*, Vol. 96, No. 2, pp. 263–276.

Mumford, E. (2006) 'The story of socio-technical design: Reflections on its successes, failures and potential', *Information Systems Journal*, Vol. 16, No. 4, pp. 317–342.

Nadler, A., Ellis, S. and Bar, I. (2003) 'To seek or not to seek: The relationship between help seeking and job performance evaluations as moderated by task-relevant expertise', *Journal of Applied Social Psychology*, Vol. 33, No. 1, pp. 91–109.

Noy, S. and Zhang, W. (2023) 'Experimental evidence on the productivity effects of generative artificial intelligence', *Science*, Vol. 381, No. 6654, pp. 187–192.

Orlikowski, W.J. (2007) 'Sociomaterial practices: Exploring technology at work', *Organization Studies*, Vol. 28, No. 9, pp. 1435–1448.

Pagnozzi, F. and Antonelli, G. (2025) 'How to seek help? A systematic literature review of help-seeking practices in organizations', *Management Review Quarterly*, pp. 1–50.

Pagnozzi, F., Renkema, M. and Antonelli, G. (2024) 'Help-seeking practices and behaviours in organizations and AI-driven chatbot: A scoping review', *PuntOorg International Journal*, Vol. 1, No. 1.

Pasmore, W., Winby, S., Mohrman, S.A. and Vanasse, R. (2019) 'Reflections: Sociotechnical systems design and organization change', *Journal of Change Management*, Vol. 19, No. 2, pp. 67–85.

Raisch, S. and Krakowski, S. (2021) 'Artificial intelligence and management: The automation–augmentation paradox', *Academy of Management Review*, Vol. 46, No. 1, pp. 192–210.

Schein, E.H. (2010) *Organizational Culture and Leadership*, 4th edn., San Francisco: Jossey-Bass.

Scott, W.R. (2013) *Institutions and Organizations: Ideas, Interests, and Identities*, Thousand Oaks: Sage Publications.

Trist, E.L. and Bamforth, K.W. (1951) 'Some social and psychological consequences of the longwall method of coal-getting', *Human Relations*, Vol. 4, No. 1, pp. 3–38.

Van Dyne, L., Ang, S. and Botero, I.C. (2003) 'Conceptualizing employee silence and employee voice as multidimensional constructs', *Journal of Management Studies*, Vol. 40, No. 6, pp. 1359–1392.

VandeWalle, D. (1997) 'Development and validation of a work domain goal orientation instrument', *Educational and Psychological Measurement*, Vol. 57, No. 6, pp. 995–1015.

Venkatesh, V., Morris, M.G., Davis, G.B. and Davis, F.D. (2003) 'User acceptance of information technology: Toward a unified view', *MIS Quarterly*, Vol. 27, No. 3, pp. 425–478.

Wanberg, C.R. and Kammeyer-Mueller, J.D. (2000) 'Predictors and outcomes of proactivity in the socialization process', *Journal of Applied Psychology*, Vol. 85, No. 3, pp. 373–385.

AI–Gamification Role Matrices in Knowledge Management: Evidence and Research Gaps

Behrooz Moradi[1], Ettore Bolisani[1], Enrico Scarso[1], Kathrin Kirchner[2], Furong Cai[1]

[1] Department of Management and Engineering, University of Padova, Vicenza, Italy

[2] Department of Engineering Technology, Technical University of Denmark, Ballerup, Denmark

Abstract

As knowledge work becomes more digital and distributed, organizations increasingly rely on two levers to sustain knowledge management (KM) performance: Artificial Intelligence (AI) to expand what KM systems can do and gamification to strengthen why people contribute and reuse knowledge. Yet prior research rarely specifies how AI and gamification are jointly organized across KM subfields, thereby limiting both theory building and design guidance. This study analyzes the AI–gamification–KM triad using a structured evidence synthesis that combines systematic searching, qualitative content analysis, and matrix-based modeling. Drawing on 44 peer-reviewed studies retrieved from four major scholarly databases with coverage through December 1, 2025, two independent coders classified each study by KM subfield, sector, outcome type, and AI and gamification role assignments. Descriptive statistics were produced in SPSS, and role matrices were constructed and visualized in R to generate a configurational evidence map linking role combinations to effect categories. The results show a clear division of labor: human involvement concentrates on learning and knowledge acquisition, while AI involvement concentrates on system-supported functions, most frequently knowledge representation, collection, and analysis. Evidence also clusters by context, with higher education dominating and outcomes primarily reported as systems and platforms. Across effect categories, the evidence base is dominated by a narrow configuration that combines engagement-oriented gamification with tailoring AI functions, while many other role combinations remain sparse or absent. These gaps are most pronounced in KM processes that require transfer, externalization, fusion, and refinement. The study contributes a framework that clarifies how AI and gamification co-configure KM outcomes and provides an evidence-based agenda for testing missing configurations across KM subfields and sectors.

Keywords: Knowledge Management, Artificial Intelligence, Gamification, Research gaps.

Paper type: Academic Research Paper.

1 Introduction

KM remains a central concern for organizations that rely on distributed digital work (Watson and Hewett, 2006; Beesley and Cooper, 2008). The rapid pace of technological change and the proliferation of information sources increase the effort needed to find, evaluate, and reuse knowledge (Tatsiopoulos and Katsanta, 2018; Fensel *et al.*, 2020; Moradi *et al.*, 2024). KM research, therefore, spans both human-centered processes and the functions of KM systems (KMSs) (Watson and Hewett, 2006; Fensel *et al.*, 2020; Moradi *et al.*, 2025).

Recent KM-related studies have examined the joint use of AI and gamification to support KM initiatives (Sánchez *et al.*, 2024; Umer *et al.*, 2024; Huang and Chuang, 2025). AI is typically studied as a set of AI-enabled functions embedded in KMSs that enable knowledge activities such as collection, validation, and dissemination of knowledge (Ipeirotis and Gabrilovich, 2014; Fensel *et al.*, 2020; Cai *et al.*, 2023). Gamification, the application of game design elements (Deterding *et al.*, 2011), is associated with improvements in knowledge acquisition and transfer (Tenorio *et al.*, 2018; Ahmed *et al.*, 2024; Youyou, Zihan and Mengfang, 2025). Therefore, studying them together matters because AI shapes KMS capability and gamification shapes user participation and engagement, and both can affect the same KM initiative.

This systematic literature review (SLR) synthesizes how AI and gamification are examined together in KM. It addresses (RQ1) how human-related and AI-related involvement is distributed across KM subfields, (RQ2) in which sectors this combination is applied and what outcome types are reported, and (RQ3) what kinds of effects are associated with different AI–gamification role configurations. The study develops a configurational evidence map that synthesizes prior work and reveals systematically underexplored AI–gamification role combinations across KM subfields.

2 Literature Review

KM research has traditionally emphasized human-driven knowledge processes through which individuals and organizations manage knowledge. Core human-related KM subfields discussed in foundational KM research include learning, knowledge acquisition, sharing, contribution, application, transfer, externalization, and creation (Nonaka, Toyama and Konno, 2000; Alavi and Leidner, 2001; Alavi and Tiwana, 2002; Watson and Hewett, 2006; Beesley and Cooper, 2008; Wang and Noe, 2010; Carleton, 2011). Within this stream, gamification is often examined as a design approach to strengthen engagement in KM activities, such as collaborative knowledge construction and refinement (Tenorio *et al.*, 2018; Ralli *et al.*, 2020; Moradi *et al.*, 2024).

In parallel, AI in KM research has emphasized system-level knowledge functions implemented within KMSs. Drawing on the Knowledge Graph lifecycle, prior work highlights functions such as knowledge collection, aggregation, validation, fusion, refinement, representation, storage, analysis, and dissemination (Fensel *et al.*, 2020). In this line of work, AI is examined as a component that can directly support KMS functions or indirectly enhance KMS performance by optimizing features that enable KM processes, including knowledge validation, refinement, and dissemination (Ipeirotis and Gabrilovich, 2014; Ralli *et al.*, 2020; Huang and Wen Chuang, 2024).

Prior reviews often examine AI, gamification, and KM separately, and they vary widely in the extent to which they cover this triad. Some reviews focus on a single AI artifact class, such as chatbots, with gamification treated as a minor facilitation, and KM coverage concentrated on knowledge acquisition in educational settings (Ma *et al.*, 2025). Related reviews treat gamification mainly as an instructional method and include only a minimal or peripheral AI role. At the same time, “knowledge” is operationalized primarily as skills or knowledge acquisition rather than as a structured set of KM subfields (Ahmed *et al.*, 2024). Other reviews integrate AI-driven personalization and gamification and include KM subfields such as learning. Still, their synthesis remains centered on learning outcomes, limiting coverage across a broader set of KM subfields (Sheejamol, Chacko and Kumar, 2025). One review maps gamification across multiple KM subfields. However, it does not include AI, so the combined evidence on AI, gamification, and KM remains incomplete and supports the need for an integrative synthesis (Moradi *et al.*, 2024).

To address these gaps, this study adopts an AI–gamification–KM lens that treats the triad as a unified configuration. This lens enables a structured synthesis of how prior studies combine these elements across KM subfields, sectors, and outcomes and forms the basis of the evidence map presented in the Results section.

3 Methodology

The study used a structured search across four databases: Scopus, Web of Science, IEEE Xplore, and the ACM Digital Library. Predefined keyword groups for gamification, KM, and AI were used to retrieve records containing terms from all three categories (Table 1). Searches covered all available years up to December 1, 2025, were limited to English-language records, and were executed on the same date.

Table 1. Search String and Keyword Categories Used in the SLR

Keywords
(“gamification” OR “game element” OR” game mechanics” OR” game dynamics” OR” Game components” OR “Gamify*"OR" gamified*"OR” Gameful”)
(“knowledge management” OR “knowledge creation” OR “knowledge acquisition” OR “knowledge storage” OR “knowledge sharing” OR “knowledge transfer” OR “knowledge application” OR “knowledge protection “OR “knowledge process“ OR “intellectual capital” OR” knowledge engineering” OR” knowledge representation” OR” knowledge mapping” OR” knowledge extraction” OR” knowledge discovery” OR” knowledge retention” OR” knowledge utilisation” OR” knowledge audit” OR “knowledge-based systems” OR” knowledge economy” OR” KM practices” OR” knowledge management practices”)
("artificial intelligence" OR "machine learning" OR "deep learning" OR "neural network" OR "natural language processing" OR "NLP" OR "expert system" OR "intelligent system*" OR "intelligent agent*" OR "automated reasoning" OR "decision support system*" OR "data mining" OR "text mining" OR "semantic system*" OR "ontolog*" OR "computational intelligence" OR "generative model*" OR "large language model*" OR "LLM*" OR "AI")

Screening was conducted in two stages: first title and abstract, then the full text. Studies were included if they were peer-reviewed journal articles or conference papers in English and contained at least one keyword from each category. Books, book chapters, theses, reports, non-peer-reviewed materials, and duplicates were excluded. Finally, 44 studies met the inclusion criteria. No backward or forward snowballing was conducted. The screening flow is presented in the PRISMA diagram (Figure 1).

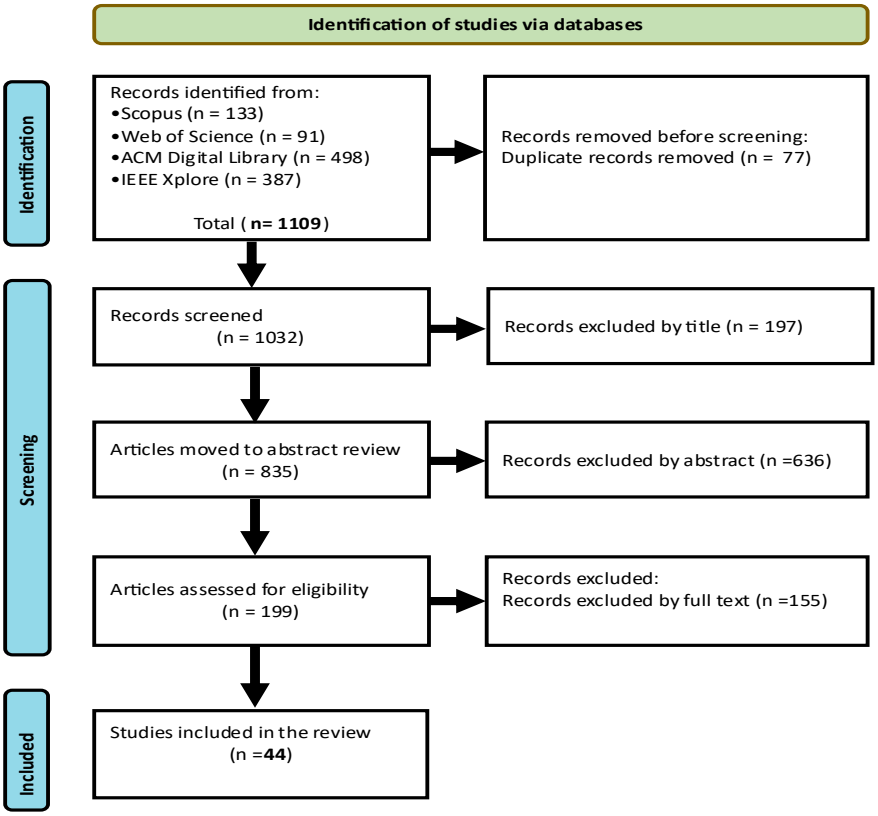


Figure 1. PRISMA Flow Diagram of Study Selection

For analysis, the included studies were examined using an SLR and qualitative content analysis through a triadic lens to assess how AI, gamification, and KM are addressed together across contexts. Study screening and coding were independently conducted by two researchers, with discrepancies resolved through consensus. Each study's KM subfield(s) and AI–gamification roles were coded, and data analysis was supported using SPSS for descriptive statistics and R for matrix construction and visualization.

Regarding the operationalization of the roles, an AI role was coded as the dominant functional contribution AI performs within the KMS. For example, AI that analyzes learner data and adapts task difficulty or generates personalized scenarios was coded as A2 (tailoring), whereas AI that optimizes system parameters through simulation was coded as A6 (system optimization). A gamification role was coded as the dominant behavioral or structural function through which game elements shape KM activities. For instance, using items to increase motivation and engagement were coded as G1, while age-appropriate items supporting knowledge acquisition were coded as G2. The detailed results are explained in the following section.

4. Result and discussion

4.1. Publication Trends and Overview of the Reviewed Study Corpus

Figure 2 shows that, although no time restrictions were applied, the earliest relevant record dates to 2013. From 2013 to 2023, output remained low, never exceeding 3 studies per year. Publication volume then increased sharply in 2024 and remained high in 2025, with 14 and 12 records, respectively. This surge reflects growing scholarly interest, likely driven by AI advancements and the need to enhance KM through motivational design. The trend confirms the triad's emergence as a cohesive research configuration and underscores the need for systematic synthesis to guide future integration and theorization.

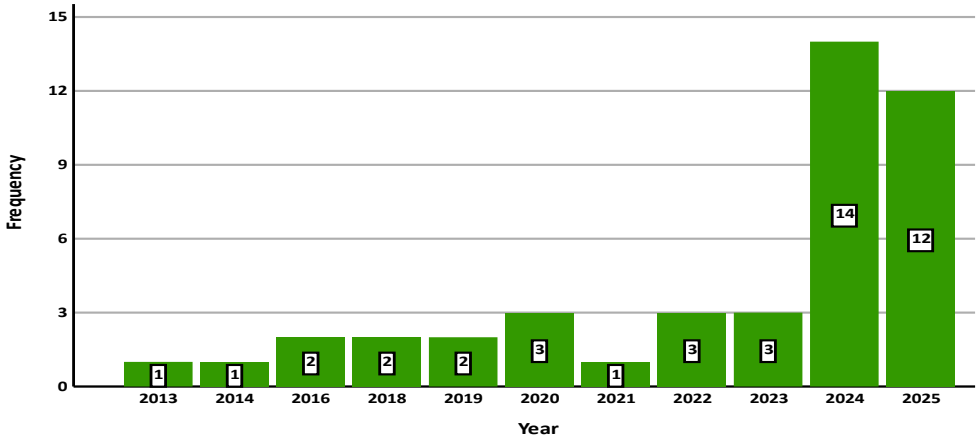


Figure 2. Annual AI–gamification–KM publications

Table 2 summarizes the coded study set and provides the data foundation for the analyses of the KM subfield, sectoral distribution, outcome types, and AI–gamification role configurations reported in Sections 4.2–4.4.

Table 2. Included studies and coding summary

Study	KM subfields	Sector	Outcome
(Plecher *et al.*, 2020)	LR	K12	Application
(Olszewski and Turek, 2016)	KC	SC	Prototype
(Shabadurai, Chua and Lim, 2025)	LR	PT	Prototype
(Dusmukhamedova *et al.*, 2024)	LR	K12	Model
(Karunasekara *et al.*, 2022)	LR	K12	Application
(Sun, 2024)	KS, KT	HE	Platform
(Lungu *et al.*, 2025)	KA	PT	System
(Huang and Chuang, 2025)	KA	HL	System
(Umer *et al.*, 2024)	LR	EO	Framework
(Mangole and Ade-ibijola, 2025)	KA	EO	Prototype
(Zirek, 2024)	KC	HE	Platform
(Quenta Leon, Romero Cacha and Navarrete Vilca, 2025)	LR	HE	Application
(Ricky *et al.*, 2024)	LR	HE	Prototype
(Ralli *et al.*, 2020)	KV, KR	SC	Platform
(Elmashhara *et al.*, 2024)	KD	EO	Model
(Fedorov, Fraternali and Pasini, 2016)	KB, KL, KG, KV, KR	SC	Application
(Tatsiopoulos and Katsanta, 2018)	KA, KS, KR, KG	HE	Model
(Missaoui, Bennani and Maalel, 2020)	LR	HE	Model
(Sorgatz, Schuldt and Gröger, 2024)	KA, KP, KR, KD	HE	Platform
(Zhao and Liu, 2024)	KC, KL, KN, KR	HE	Platform
(Sánchez *et al.*, 2024)	LR	HE	Platform
(Redondo *et al.*, 2025)	LR	HE	Application
(You, Chen and Zhao, 2025)	LR	HE	Prototype
(Singh and Singh, 2025)	LR	HE	System
(Ho and Chien-Sing, 2024)	LR	HE	Prototype
(Narayanan and Kumaravel, 2023)	LR	HE	System
(Mishra *et al.*, 2025)	LR, KR	HE	Platform
(Aleksandrovich *et al.*, 2024)	KA, KN	HE	Analysis
(Dermeval *et al.*, 2019)	LR, KR	HE	Model
(Youyou, Zihan and Mengfang, 2025)	TR	HE	Model
(Huang and Wen Chuang, 2024)	KA, KD	HL	System
(Vu *et al.*, 2022)	KL, KF, KR, KT	HL	Platform
(Bölz *et al.*, 2024)	KA	HL	System
(Sevastjanova *et al.*, 2024)	KE, KB, KL, KN	CKS	Framework
(Kumar *et al.*, 2022)	KA, KN	K12	System
(Tenorio *et al.*, 2018)	KA	K12	Application
(Anunpattana, Khalid and Iida, 2023)	LR, KR	K12	Model
(Stannett, Legg and Sarjant, 2013)	KS, KC, KR	CKS	System

(Purna Kumari *et al.*, 2025)	KP, RF	PT	System
(Cai *et al.*, 2023)	LR, KD	SC	System
(Ipeirotis and Gabrilovich, 2014)	KL, KV	CKS	System
(Martin, Torres and Fernandez, 2021)	KC, KL, KG, KN	CKS	Approach
(Khelifi *et al.*, 2024)	KS	PT	Platform
(Sevastjanova *et al.*, 2019)	KC, KS	CKS	Model

Abbreviations: Learning (LR), Knowledge creation (KC), Knowledge sharing (KS), Knowledge storage (KT), Knowledge acquisition (KA), Knowledge representation (KR), Knowledge aggregation (KG), Knowledge dissemination (KD), Knowledge application (KP), Knowledge collection (KL), Knowledge analysis (KN), Knowledge transfer (TR), Knowledge validation (KV), Knowledge contribution (KB), Knowledge externalization (KE), Knowledge fusion (KF), Knowledge refinement (RF).
K–12 Education (K12), Higher Education (HE), Computational Knowledge Systems (CKS), Health (HL), Professional Training (PT), Enterprise Operations (EO), Societal Contexts (SC).

4.2. Human and AI Involvement Across KM Subfields (RQ1)

Figures 3 and 4 visualize the resulting distributions for RQ1 and RQ2. Figure 3 uses color coding to distinguish human-related (green) and AI-related (blue) subfields. Human involvement is concentrated in learning (40.0%) and knowledge acquisition (22.2%), with smaller shares in knowledge creation and knowledge sharing. At the lower end, knowledge externalization and knowledge transfer are least represented (2.2% each). Regarding AI involvement, knowledge representation is most frequent (28.6%), followed by knowledge collection and knowledge analysis, while knowledge fusion and knowledge refinement are least frequent (2.9% each). This distribution indicates that the literature most often places humans at the center of KM subfields that depend on learning and interpersonal exchange, where meaning and context are critical.

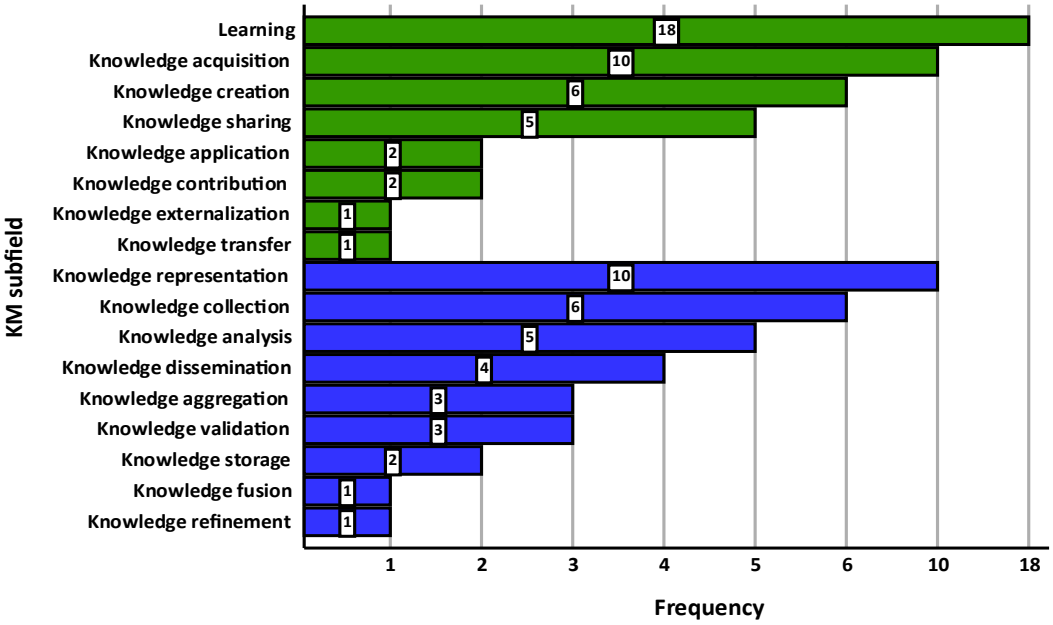


Figure 3. Human and AI involvement across KM subfields
Note: Studies may contribute to multiple subfields

AI involvement clusters in subfields that structure and process knowledge inside KMSs, especially representation, collection, and analysis. The pattern suggests that current studies more often treat AI as infrastructure that organizes and delivers knowledge rather than as a direct substitute for human knowledge work. Knowledge transfer, externalization, fusion, and refinement appear only rarely in the reviewed studies. As a result, the evidence is limited on how human and AI involvement is reported and characterized for these KM processes across sectors.

4.3. Sectoral Outcome Patterns in KM Studies (RQ2)

Figure 4 summarizes the reviewed studies by sector and outcome type, showing a pronounced concentration in three domains. Higher education (HE) accounts for 18 studies, followed by K–12 education (6) and computational knowledge systems (5). The remaining sectors, including societal contexts, professional training, and health, are evenly represented, each contributing 4 studies, while enterprise operations contribute 3

studies. Regarding outcomes, *systems* are the most frequently reported artifact, appearing in 11 studies, followed by *platforms* in 9 and *models* in 8. Other outcomes, such as prototypes and applications, occur less frequently, while frameworks, approaches, and analytical studies are comparatively limited.

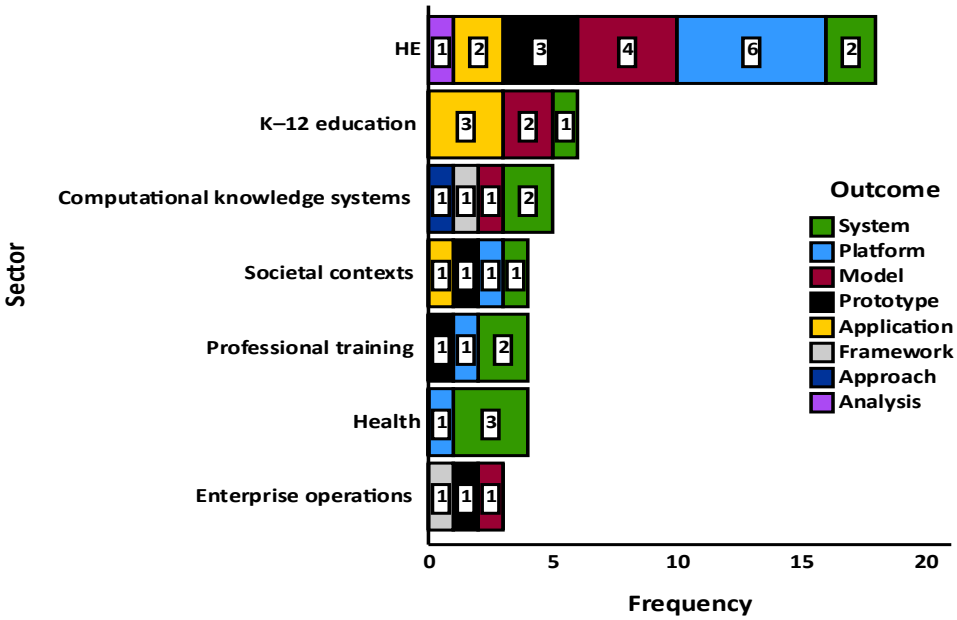


Figure 4. Sector and outcome distribution of studies

4.4. AI–Gamification Role Configurations and Associated Effects (RQ3)

Based on a cross-study synthesis, Figure 5 synthesizes the reviewed evidence by illustrating how AI roles (A1–A7) and gamification roles (G1–G7) are combined across studies addressing knowledge-related contexts to produce distinct effect categories (F1–F7). Across the seven matrices, a clear concentration emerges around G1 (82) and A2 (42), with their co-occurrence (G1–A2; 27) as the most frequent configuration. This dominance indicates that tailoring AI functions (recommendation, adaptation, and personalization) are predominantly used to steer engagement-oriented gamification, while most other AI–gamification role combinations remain sparsely examined. Building on this overall concentration, the following patterns differentiate how specific AI–gamification role configurations align with distinct effect categories.

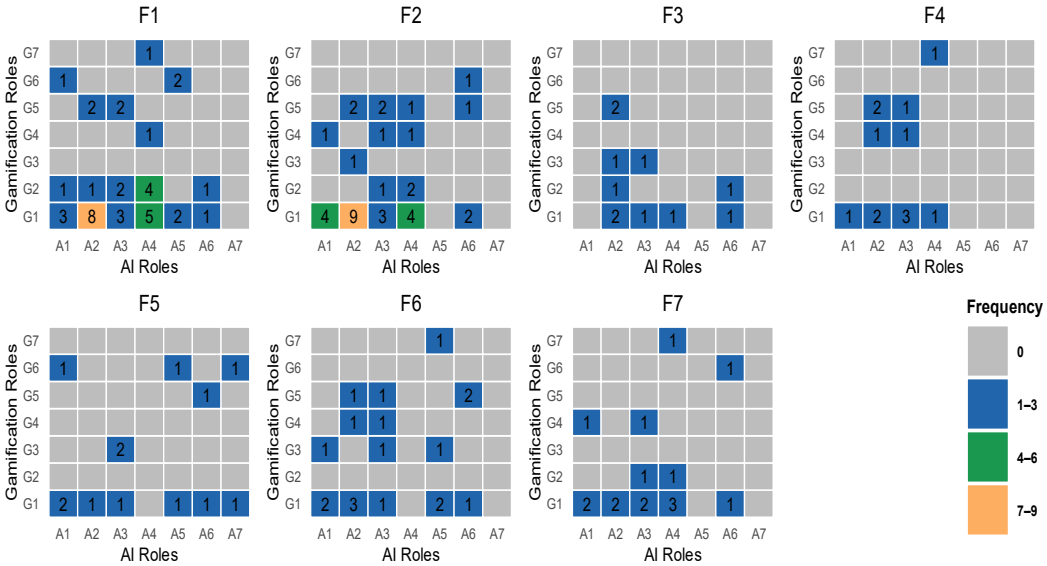
Engagement-centered, low-complexity configurations characterize F1 and F2, where effects are primarily supported by G1 combined with A2 and A4. At the same time, F3 reflects a weaker extension of this pattern with a modest presence of A6, indicating reliance on engagement and tailoring approaches with limited system-driven reinforcement.

Balanced dual-role configurations emerge in F4, where G1 is combined equally with A2 and A3, suggesting that task-related effects are supported through the joint contribution of tailored guidance and information provision rather than a single dominant AI role.

Distributed, multi-role configurations are evident in F5 and F6, where G1 co-occurs with multiple AI roles (A1, A3, A5, A6), indicating that creation- and system-level effects are associated with broader AI role combinations.

Interaction-focused configurations characterize F7, where G1 is combined with A4 and complemented by A3 and A1, highlighting the importance of interactive and informational support for satisfaction-related effects.

Overall, the findings show that AI–gamification effects in knowledge-related contexts depend on how specific roles are configured, rather than on AI or gamification alone. Unpopulated and sparsely populated matrix cells highlight evidence-based research gaps that point to unexplored AI–gamification role configurations.



Code	AI Roles	Description
A1	Detect / Analyze / Assess	Recognize patterns, classify inputs, and analytically assess data without prescribing actions.
A2	Recommend / Adaptation/ Personalization	Provide personalized recommendations or adapt content, difficulty, or pathways based on user data.
A3	Content Generating	Automatically generate content such as text, quizzes, explanations, or questions.
A4	Conversational Support	Deliver dialog-based interaction through chatbots or virtual tutors to support users.
A5	Knowledge Representation & Reasoning	Structure, represent, and reason over knowledge using ontologies, rules, or knowledge bases.
A6	Simulate / Optimize / Control	Simulate environments, manage agents, or optimize system parameters and control processes.
A7	Extract / Annotate / Validate	Extract metadata, propose annotations, or support human validation and curation tasks.

Code	Gamification Roles	Description
G1	Motivation & Engagement	Increase general motivation, engagement, persistence, and enjoyment.
G2	Knowledge acquisition and Learning	Support knowledge acquisition, learning outcomes, and skill development.
G3	Behavior Change	Encourage long-term behavior change (e.g., health, sustainability, compliance).
G4	Progression & Competition Structuring	Organize progression through levels, challenges, goals, and competitive structures.
G5	Effort & Quality Recognition	Recognize effort or contribution quality using points, badges, rewards, or achievements.
G6	Knowledge Contribution & Collaboration	Drive social participation, crowdsourcing, and collaborative knowledge contribution.
G7	Decision & Pathway Guidance	Guide user decisions and learning paths through nudging, branching, or adaptive choice structures.

Code	Findings	Description
F1	Knowledge Acquiring & Transfer	AI-enabled personalization and collaborative tools were associated with stronger knowledge acquisition and transfer.
F2	Participation Intensity & Use	Interactive system design and engagement mechanisms were linked to higher participation and sustained use.
F3	Sustained Behavior Adoption	Personalized training and persuasive gamification supported continued behavioral change and compliance.
F4	Decision & Task Performance	Prediction, recommendation, and personalization tools were associated with better decisions and task performance.
F5	Collaborative Knowledge Creation and Review	Collaboration, crowdsourcing, AI evaluation, and contribution validation enhanced knowledge creation and contribution quality.
F6	System Effectiveness & Efficiency	AI optimization, ontologies, recommendations, and automation contributed to greater system effectiveness and efficiency.
F7	User Satisfaction & Perceived KMS Value	Engaging, explainable, and user-friendly systems were linked to higher satisfaction and perceived KMS value.

Figure 5. AI–gamification role configurations and associated effects

5 Conclusion

This study examines how AI and gamification are jointly applied in KM using a configurational perspective. Through an SLR and qualitative content analysis, it synthesizes KM subfields, sectors, outcomes, and AI–gamification roles, supported by analysis in SPSS and matrix-based modeling in R.

The findings show that human involvement concentrates on learning- and acquisition-oriented KM subfields, while AI involvement is most prominent in system-supported KM functions such as representation, collection, and analysis, with studies predominantly situated in higher education and reporting system- and platform-level outcomes. Engagement-oriented gamification combined with tailoring AI functions emerges as the dominant configuration.

Theoretically, the study advances KM research by conceptualizing AI and gamification as configurable roles whose alignment explains variation in KM effects. This lens clarifies how human and system roles interact across KM subfields and shape distinct effect categories.

Methodologically, the study contributes a replicable, role-based configurational approach, operationalized through matrix-based evidence mapping, which enables systematic cross-study comparison of AI–gamification role combinations and their effects. This approach supports both synthesis and explicit identification of underexplored configurations.

This study offers practical implications for designing AI-enabled gamified KMS. The evidence indicates that personalization-oriented AI, such as adaptive feedback or content recommendations, commonly complements engagement-focused gamification when the goal is to increase user participation. For more complex KM objectives, such as structured task performance or knowledge creation, broader combinations of analytical, optimization, and guidance-oriented AI capabilities appear more suitable than engagement-centered designs alone. Thus, the matrices support informed managerial decisions when configuring AI and gamification for specific knowledge goals.

Finally, the presence of sparsely populated and empty matrix cells reveals evidence-based research gaps, pointing to AI–gamification role combinations that remain largely unexplored. Future research should empirically examine these missing configurations across KM subfields and sectors to assess their effects and boundary conditions. This study is limited by its reliance on peer-reviewed, English-language publications retrieved from selected academic databases, which may exclude relevant work published elsewhere or in other languages.

6 Acknowledgments

The study that led to these results received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No. 101034319 and from the European Union's NextGenerationEU.

7 Ethics Declaration

This study relies on an SLR and content analysis of published studies and, therefore, did not require ethical approval.

8 AI Declaration

AI tools (ChatGPT 5.3 and Grammarly) were used solely to improve grammar and sentence structure.

References

Ahmed, A. *et al.* (2024) “How universities teach cybersecurity courses online: a systematic literature review,” *Frontiers in Computer Science*, 6, p. 1499490.

Alavi, M. and Leidner, D.E. (2001) “Knowledge Management and Knowledge Management Systems: Conceptual Foundations And Research Issues,” *MIS Quarterly*, 25(1), pp. 107–136.

Alavi, M. and Tiwana, A. (2002) “Knowledge integration in virtual teams: The potential role of KMS,” *Journal of the American Society for Information Science and Technology*, 53(12), pp. 1029–1037.

Aleksandrovich, S.I. *et al.* (2024) “Transformative applications in biology education: A case study on the efficacy of adaptive learning with numerical insights,” *Caspian Journal of Environmental Sciences*, 22(2), pp. 395–408.

Anunpattana, P., Khalid, M.N.A. and Iida, H. (2023) “Objectivity and Subjectivity in Variation of Multiple Choice Questions: Linking the Theoretical Concepts Using Motion in Mind,” *IEEE Access*, 11, pp. 35371–35397.

Beesley, L.G.A. and Cooper, C. (2008) “Defining knowledge management (KM) activities: towards consensus,” *Journal of Knowledge Management*, 12(3), pp. 48–62.

Bölz, F. *et al.* (2024) "Food KBQA Recommender: A Knowledge Base Question Answering System Powered Meal Plan Recommender," *IEEE/WIC International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT)*. Bangkok, Thailand: IEEE, pp. 86–93.

Cai, X. *et al.* (2023) "Research and Design of Beachcombing Service System Based on Marine Ecotourism Concept." *16th International Symposium on Computational Intelligence and Design (ISCID)*, Hangzhou, China: IEEE, pp. 218–221.

Carleton, K. (2011) "How to Motivate and Retain Knowledge Workers in Organizations: A Review of the Literature," *International Journal of Management*, 28(2), pp. 459–468.

Dermeval, D. *et al.* (2019) "GaTO: An Ontological Model to Apply Gamification in Intelligent Tutoring Systems," *Frontiers in Artificial Intelligence*, 2, p. 13.

Deterding, S. *et al.* (2011) "From game design elements to gamefulness: defining 'gamification,'" *Proceedings of the 15th international academic MindTrek conference: Envisioning future media environments*, pp. 9–15.

Dusmukhamedova, S. *et al.* (2024) "A Novel Based Design of AI kind of Models for Gami-Fication in Enhanced Learning type of Environment." *4th International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*, Greater Noida, India: IEEE, pp. 584–588.

Elmashhara, M.G. *et al.* (2024) "How gamifying AI shapes customer motivation, engagement, and purchase behavior," *Psychology & Marketing*, 41(1), pp. 134–150.

Fedorov, R., Fraternali, P. and Pasini, C. (2016) "SnowWatch: A Multi-modal Citizen Science Application," in A. Bozzon, P. Cudre-Maroux, and C. Pautasso (eds.) *Web Engineering*. Cham: Springer International Publishing, pp. 538–541.

Fensel, D. *et al.* (2020) *Knowledge Graphs: Methodology, Tools and Selected Use Cases*. Cham: Springer International Publishing.

Ho, H.X.A. and Chien-Sing, L. (2024) "Effectiveness of the design of a gamified Voice-Assisted Chatbot system in aiding modelling, and in personalizing guidance for learning basic OOP," *Proceedings of International Conference on Computational Thinking Education*. The Education University of Hong Kong.

Huang, H. and Wen Chuang, H. (2024) "Revolutionizing mHealth Interaction with a Gamified Chatbot: An OMO Strategy Approach." *20th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, Paris, France: IEEE, pp. 283–288.

Huang, H.C. and Chuang, H.W. (2025) "A Pilot Study on AI-Powered Gamified Chatbot with OMO Strategy for Enhancing Parental Nutrition Knowledge," *Digital*, 5(2), p. 13.

Ipeirotis, P.G. and Gabrilovich, E. (2014) "Quizz: targeted crowdsourcing with a billion (potential) users," *Proceedings of the 23rd International Conference on World Wide Web*. New York, NY, USA: Association for Computing Machinery (WWW '14), pp. 143–154.

Karunasekara, K.K.R. *et al.* (2022) "A Mobile Gamified Application to Improve Mathematics Skills of Students from Age 9 to 11 Years." *10th Region 10 Humanitarian Technology Conference (R10-HTC)*, Hyderabad, India: IEEE, pp. 136–141.

Khelifi, J. *et al.* (2024) "GitRev: An LLM-Based Gamification Framework for Modern Code Review Activities." *International Conference on Source Code Analysis and Manipulation (SCAM)*, Flagstaff, AZ, USA: IEEE, pp. 235–241.

Kumar, N.S. *et al.* (2022) "Identification and Addressal of Knowledge Gaps in Students." *3rd International Conference for Emerging Technology (INCET)*, Belgaum, India: IEEE, pp. 1–6.

Lungu, N. *et al.* (2025) "Adaptive Gamified Cybersecurity Training for Enhanced User Engagement." *5th International Conference on Soft Computing for Security Applications (ICSCSA)*, Salem, India: IEEE, pp. 124–129.

Ma, Wenjuan *et al.* (2025) "The who, why, and how of AI-based chatbots for learning and teaching in higher education: A systematic review," *Education and Information Technologies*, 30(6), pp. 7781–7805.

Mangole, P.S. and Ade-ibijola, A. (2025) "A Personalised Learning Platform with Text Generation, Games, and Chatbots for Mastering Business Administration Concepts," *Journal of Learning and Teaching in Digital Age*, 10(2), pp. 170–186.

Martin, J., Torres, D. and Fernandez, A. (2021) "Optimizing a Gamified Design Through Reinforcement Learning - a Case Study in Stack Overflow," in M. Naiouf et al. (eds.) *Cloud Computing, Big Data & Emerging Topics*. Cham: Springer International Publishing, pp. 89–103.

Mishra, R. *et al.* (2025) "InvestED: A Smart Financial Education Platform Using LSTM & Game Mechanics." *5th International Conference on Intelligent Technologies (CONIT)*, HUBBALI, India: IEEE, pp. 1–8.

Missaoui, S., Bennani, S. and Maalel, A. (2020) "Towards an Ontology for Representing a Student's Profile in Adaptive Gamified Learning System." *International Conference on Knowledge Engineering and Ontology Development*, Budapest, Hungary: SCITEPRESS, pp. 149–156.

Moradi, B. *et al.* (2024) "Gamification and its Relevance to Knowledge Management: A Systematic Literature Review." *19TH INTERNATIONAL FORUM ON KNOWLEDGE ASSET DYNAMICS, IFKAD 2024*.

Moradi, B. *et al.* (2025) "Towards a Unified Framework for Knowledge Worker Roles: A Systematic Literature Review," *Proceedings of the 20th IFKAD Conference*. University of Padua Research Archive, pp. 1385–1394.

Narayanan, K.S. and Kumaravel, A. (2023) "Gamification and XAI Tutoring Structure using Mamdani Fuzzy Inference Systems." *9th International Conference on Smart Structures and Systems (ICSSS)*, CHENNAI, India: IEEE, pp. 1–5.

Nonaka, I., Toyama, R. and Konno, N. (2000) "SECI, Ba and Leadership: a Unified Model of Dynamic Knowledge Creation," *Long range planning*, 33(1), pp. 5–34.

Olszewski, R. and Turek, A. (2016) "Application of the Spatial Data Mining Methodology and Gamification for the Optimisation of Solving the Transport Issues of the 'Varsovian Mordor,'" in Y. Tan and Y. Shi (eds.) *Data Mining and Big Data*. Cham: Springer International Publishing, pp. 103–114.

Plecher, D.A. *et al.* (2020) “ARsinoë - Learning Egyptian Hieroglyphs with Augmented Reality and Machine Learning.” *International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, Recife, Brazil: IEEE, pp. 326–332.

Purna Kumari, P.V. *et al.* (2025) “Gamification and Blockchain: Innovative Approaches to Employee Motivation.” *6th International Conference for Emerging Technology (INCET)*, BELGAUM, India: IEEE, pp. 1–5.

Quenta Leon, L.A., Romero Cacha, R.V. and Navarrete Vilca, E.J. (2025) “A Cognitive AI Model Using Deep & Cross Network for Personalized Gamified Recommendations to Promote Healthy Eating Decisions Among University Students.” *3rd Cognitive Models and Artificial Intelligence Conference (AICCONF)*, Prague, Czech Republic: IEEE, pp. 1–7.

Ralli, M. *et al.* (2020) “CrowdHeritage: Improving the quality of Cultural Heritage through crowdsourcing methods.” *15th International Workshop on Semantic and Social Media Adaptation and Personalization (SMAP)*, Zakynthos, Greece: IEEE, pp. 1–6.

Redondo, J.I. *et al.* (2025) “Choose Your Own Adventure: Using Twine for Gamified Interactive Learning in Veterinary Anaesthesia,” *Veterinary Sciences*, 12(2), p. 156.

Ricky, S.M.S.I. *et al.* (2024) “Applied Instructional Design in Programee to Enhance Students E-learning Experience: Programee - The Future of E-Learning App.” *3rd International Conference on Computing Advancements*, Dhaka Bangladesh: ACM, pp. 239–246.

Sánchez, J.B. *et al.* (2024) “Gamified and Artificial Intelligence-Assisted Knowledge Base in a Social Platform for Self-Learning in Higher Education.” *International Symposium on Accreditation of Engineering and Computing Education (ICACIT)*, Bogota, Colombia: IEEE, pp. 1–6.

Sevastjanova, R. *et al.* (2019) “Shall we play? – Extending the Visual Analytics Design Space through Gameful Design Concepts.” *Workshop on Machine Learning from User Interaction for Visualization and Analytics (MLUI)*, Vancouver, BC, Canada: IEEE, pp. 1–9.

Sevastjanova, R. *et al.* (2024) “Personalized Language Model Selection Through Gamified Elicitation of Contrastive Concept Preferences,” *IEEE Transactions on Visualization and Computer Graphics*, 30(8), pp. 5449–5465.

Shabadurai, Y., Chua, F.-F. and Lim, T.-Y. (2025) “A Reinforcement Learning-Based Dynamic Adaptive Gamification (DAG) Method for Online Employee Training.” *14th International Conference on Information Technology in Asia (CITA)*, Kota Samarahan, Malaysia: IEEE, pp. 66–71.

Sheejamol, P.T., Chacko, A.M. and Kumar, S.D.M. (2025) “Beyond the One-Size-Fits-All: A Systematic Review of Personalized and Gamified e-Learning for Neurodivergent Learners,” *Electronic Journal of e-Learning*, 23(3), pp. 101–119.

Singh, K.D. and Singh, P.D. (2025) “Edge-AI Driven Gamification in Engineering Education for Improving Student Engagement,” *Computer Applications in Engineering Education*, 33(5), p. e70076.

Sorgatz, A., Schuldt, J. and Gröger, S. (2024) “Virtual-tutoring-supported teaching for Geometrical Product Specification,” *Procedia CIRP*, 129, pp. 216–221.

Stannett, M., Legg, C. and Sarjant, S. (2013) “Massive Ontology Interface,” *Proceedings of the 14th Annual ACM SIGCHI_NZ Conference on Computer-Human Interaction*. New York, NY, USA: Association for Computing Machinery (CHINZ ’13).

Sun, G. (2024) “Blockchain-Supported Platform for Collaborative Livestreaming of STEM Courses.” *4th International Conference on Electronic Information Engineering and Computer (EIECT)*, Shenzhen, China: IEEE, pp. 528–533.

Tatsiopoulos, C. and Katsanta, D.-M. (2018) “On ontologies and knowledge associations in gamified Environments,” *2018 9th International Conference on Information, Intelligence, Systems and Applications (IISA)*. IEEE, pp. 1–7.

Tenorio, M.M. *et al.* (2018) “Céos: A Collaborative Web-Based Application for Improving Teaching-Learning Strategies,” in M.E. Auer and T. Tsiatsos (eds.) *Interactive Mobile Communication Technologies and Learning*. Cham: Springer International Publishing, pp. 107–114.

Umer, W. *et al.* (2024) “A study on the Impact of how AI-powered Green Development Influence Employee Green Behavior within an Organization.” *Horizons of Information Technology and Engineering*, Lahore, Pakistan: IEEE, pp. 1–5.

Vu, B. *et al.* (2022) “Towards Continuous Professional Monitoring of Health Status Based on Energetic Balancing.” *International Workshop on Sport, Technology and Research (STAR)*, Trento - Cavalese, Italy: IEEE, pp. 72–77.

Wang, S. and Noe, R.A. (2010) “Knowledge sharing: A review and directions for future research,” *Human Resource Management Review*, 20(2), pp. 115–131.

Watson, S. and Hewett, K. (2006) “A Multi-Theoretical Model of Knowledge Transfer in Organizations: Determinants of Knowledge Contribution and Knowledge Reuse,” *Journal of Management Studies*, 43(2), pp. 141–173.

You, Y., Chen, H.W. and Zhao, J. (2025) “Enhancing AI Explainability for Non-technical Users with LLM-Driven Narrative Gamification,” *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*. New York, NY, USA: ACM (CHI EA ’25).

Youyou, Zihan, C. and Mengfang, H. (2025) “Interdisciplinary Integration of Knowledge Construction Based on the ARCS Motivational Model in AI-Driven Gamified Learning Environments.” *5th International Conference on Artificial Intelligence and Education (ICAIE)*, Suzhou, China: IEEE, pp. 597–603.

Zhao, W. and Liu, X. (2024) “NxTEL: Next Generation Technology-Enhanced Learning Platform for Computer Science Education.” *Integrated STEM Education Conference (ISEC)*, Princeton, NJ, USA: IEEE, pp. 1–8.

Zirek, M. (2024) “A Novel Approach to University Startup Incubators by Using AI Powered Simulation and Gamification.” *International Conference on Computing, Networking, Telecommunications & Engineering Sciences Applications (CoNTESA)*, Tirana, Albania: IEEE, pp. 1–6.

Decision-Making Approaches and Entrepreneurs' Pivot Decisions? An RCT Experiment

Micol Di Vita, Andrea Panelli
Department of Management and Production Engineering, Politecnico di Torino, Turin, Italy

Abstract

Do decision-making approaches affect entrepreneurs' pivot decisions on the Business Model Canvas? To address our research question, we investigated how decision-making approaches, namely the scientific method (Camuffo et al., 2024) and the effectuation method (Sarasvathy, 2008), affect entrepreneurs' pivot decisions. In particular, we analysed how entrepreneurs tend to change blocks on the left or right of the Business Model Canvas BMC (Osterwalder, 2004), depending on the methods they are approaching. In order to achieve our goal, we based our research on entrepreneurial learning (Cohen et al., 2019; McDonald and Eisenhardt, 2020; Pillai et al. 2020), included in entrepreneurial theories (Ott and Hannah, 2024) and we applied it to the scientific method, according to the theory-and-evidence-based approach (Agarwal et al., 2024), and to Saras Sarasvathy's theory of effectuation (Sarasvathy, 2008), related to the evidence-based approach. By building on previous literature, little is known about which specifically BMC's blocks entrepreneurs adopting these approaches tend to change, in particular during entrepreneurial training programs. In this vein, we collected evidence via a randomized control trial (RCT) experiment with 308 Italian early-stage startups randomly assigned to "treatment" (scientific, effectuation) and "control" group (trial-and-error), involved in a pre-acceleration program. We observed early-stage startups throughout and post-training and we compared differences concerning which side of the BMC entrepreneurs changed and how many times they made pivots to modify the BMC. We find that entrepreneurs using a scientific method are more likely to make one pivot both on the left and on the right side of the BMC compared to the control group. On the other hand, entrepreneurs using an effectuation method tend to make at least one pivot only on the right side compared to the control group. Furthermore, considering the cumulate number of pivots, we found that entrepreneurs who learn a scientific method made more pivots on the left side of the BMC. On the other hand, entrepreneurs approaching the effectuation method made more pivots on the right side of the BMC. These findings demonstrate that the scientific method and effectuation method can shape pivot decisions differently.

Keywords: decision-making, scientific approach, pivot, randomized control trial, business model canvas

Paper type: Academic Research Paper

1 Introduction

At the origin of a startup, entrepreneurs usually build a business model to implement their entrepreneurial idea. This model includes nine blocks (value proposition, customer segments, customer relationships, channels, revenue streams, key activities, key resources, key partners and cost structure) necessary for the survival of the startup under extreme uncertainty. To address this uncertainty, previous research has proposed Business Model Canvas BMC (Osterwalder, 2004), as a tool to visualize the components necessary for entrepreneurial idea. Entrepreneurs typically represent the BMC with blocks on the left side that manage infrastructure (key partners, key activities, key structure, cost structure) and blocks on the right side related to value creation (value proposition, customer relationships, customer segments, channels, revenue streams). For instance, during early-stage phase entrepreneurs need to verify frequently their business model to address problems or strategic additions. In this vein entrepreneurs tend to make a pivot, which represents a strategic action to reallocate activities, resources and attention through an accumulated series of decisions by reorienting the startup's direction (Kirtley & O'Mahony, 2020). Moreover, pivots are knowledge-adaptation processes that can contribute to the sustainability of the business model development. Based on previous literature, startups create an iterate process to validate assumptions with high frequency of changes across multiple components of the BMC (Bellavitis et al., 2024). They aim to respond to market feedback and align their innovations with market needs. In particular, entrepreneurs tend to change the block of value proposition and customer segments because these are the levers for finding product–market; on the other hand, key blocks remain below-average developed (Slavik, 2019). This perspective is consistent with trial-and-error epitomized by Lean Startup methodologies (Ries, 2011) but it does not specify how many pivots might be needed nor how to halt the decision-making process (Sarasvathy, 2024). For this reason, recent research has proposed various interventions to help entrepreneurs' pivot decisions, typically provided in the form of entrepreneurial training (ET) programs (Alexandria et al., 2014; Bacigalupo et al., 2016; Morris et al., 2023). ET programs aim to enhance entrepreneurs' decision-making processes by helping them collect information to refine their BMC and better understand

whether to pivot. Furthermore, there is an existing research stream showing that entrepreneurs who learn specific decision-making methods through ET programs perform differently from those who do not (Camuffo et al., 2024). In this context, researchers indicate that entrepreneurs trained with the scientific method are more likely to make fewer strategic pivots, showing how they seem thrifty with resources and more strategic-oriented. In addition, another approach that focuses on resources' availability is Sarasvathy's effectuation method, according to which entrepreneurial process moves forward only through commitments that make the next step affordable loss (Sarasvathy, 2024). These outcomes occur because these entrepreneurs, by leveraging specific decision-making frameworks, collect different types of information and feedback that ultimately shape their pivot decisions. Nevertheless, we lack evidence on which specifically BMC's blocks entrepreneurs tend to change when they adopt the scientific method or the effectuation method during ET programs. This gap is particularly relevant for both scholars and practitioners, as different training programs can prompt entrepreneurs to change different BMC's blocks, choosing to focus on a specific side of the BMC. In that sense, this research aims to investigate how entrepreneurs tend to change blocks on the left or right of the BMC depending on the methods they are approaching. Given these premises, we address the following research question *"Do decision-making approaches affect entrepreneurs' pivot decisions on the Business Model Canvas?"*. We based our research on entrepreneurial learning (Cohen et al., 2019; McDonald and Eisenhardt, 2020; Pillai et al. 2020), included in entrepreneurial theories (Ott and Hannah, 2024) and we applied it to the scientific method, according to the theory-and-evidence-based approach (Agarwal et al., 2024), and to Saras Sarasvathy's theory of effectuation (Sarasvathy, 2008), related to the evidence-based approach. In order to achieve our goal, we leveraged data collected through a randomized control trial (RCT) embedded in a pre-acceleration program. We randomly assigned 308 Italian entrepreneurs to three different groups, each exposed to a distinct entrepreneurial training. One group received training in a scientific decision-making method (Camuffo et al., 2020), another group in effectuation (Sarasvathy, 2008), and the last group in a heuristic approach emphasizing trial and error. We observed early-stage startups throughout and post-training and we compared differences concerning which side of the BMC entrepreneurs changed and how many times they made pivots to modify the BMC. Our econometric analyses show that entrepreneurs who learn a scientific method seem to adjust the left side of the BMC related to infrastructure after the value creation side is validated. On the other hand, entrepreneurs adopting an effectuation method prefer to look for blocks on the right side of the BMC, in particular related to customers. As results, both scientific and effectuation methods impact on entrepreneurs' pivot decisions, showing which side of the BMC they prefer to change.

2 Theoretical Background and Hypothesis Development

Previous literature has focused on entrepreneurial learning (Cohen et al., 2019; McDonald and Eisenhardt, 2020; Pillai et al. 2020), included in entrepreneurial theories (Ott and Hannah, 2024), to develop mental models able to organize and capture a business opportunity. In this context, entrepreneurs learn how to develop theories that can be combined to "fragments" of experience and observation. In this way, entrepreneurs can extend beyond existing experiences and observations to create cohesive understanding about how to act (Ott and Hannah, 2024). This combination is achieved through experimentation that works and illustrates its intersections with other learning processes and commitment. Entrepreneurs combine action and cognition by using business-model templates to guide experiments. They recognize the role of Business Model Canvas as cognitive schema, providing broad guidance on developing early prototypes that advance learning and create value (McDonald and Eisenhardt, 2020). Such experiments are the key to success for those startups that have associated learning to economical experiments to make strategic pivots (Pillai et al. 2020). For this reason, entrepreneurial training programs (ET) for early-stage startups play a vital role in supporting entrepreneurs combining theory and experimentation. These programs seek to provide entrepreneurs with decision-making abilities, enabling them to shape a new business and understand whenever make pivotal decisions. More precisely entrepreneurs, who learn specific decision-making approach through ETs have a better understanding of the value of their idea, thus the number of pivots made (Camuffo et al., 2020). For instance, research on programs based on teaching a scientific method to decision-making has shown that entrepreneurs who learn this approach are more likely to make fewer changes to their value proposition and customer segments (Camuffo et al., 2024). On the other hand, entrepreneurs who learn effectuation method are explicitly open to changing their goals and they can assent to a wider variety of products and services sooner (Sarasvathy, 2024). Despite these premises, little is known about which specifically BMC's blocks entrepreneurs tend to change when they adopt the scientific method or the effectuation method. In particular, which side of the BMC entrepreneurs prefer to change during ET programs.

Building on these studies, we argue that different decision-making methods taught in ET can push entrepreneurs to pivot on different parts of their BMC. Due scientific entrepreneurs tend to make fewer pivots to the value creation side, we propose that they will be more inclined to explore the left side of the BMC related to infrastructure than other entrepreneurs; on the other hand, due entrepreneurs approaching the effectuation method prefer to satisfy customers interest, we expect that they will prefer to focus more on pivoting the right side of the BMC related to value creation compared to other entrepreneurs.

Accordingly, we propose the following hypothesis:

H: *Different decision-making methods push entrepreneurs to pivot on different parts of their BMC.*

3 Methodology and Data

3.1 Methodology

We collected evidence via a randomized control trial (RCT) in which 308 early-stage startups from Italy involved in a pre-acceleration program were randomly assigned to "treatment" (scientific, effectuation) and "control" group, similarly to previous studies (Camuffo et al., 2024). At the beginning of the experiment, we randomly assigned entrepreneurs to two treatment groups (being taught how to develop a business model using the scientific method; and being taught how to develop a business model using the effectuation method). We also assigned entrepreneurs to a third control group where they approached trial-and-error and they have to develop a business model using heuristics.

The pre-acceleration program was delivered online (due to Covid-19 lockdowns) and it provided training to early-stage startups for a short period of time (four months and a half). After the call for application, our initial samples include 308 early-stage startups. Using randomization, we divided the 308 startups between the two treatment groups and the control group. We checked that each treatment group (102 startups for the scientific method and 105 startups for the effectuation method) and the control group (101 startups) were balanced on several key covariates that might affect the absorption of the treatments and subsequent outcomes. Differences between the treatment groups and control are small in magnitude, and there are no significant differences between the three groups. The pre-acceleration program lasted from October 2020 to February 2021. The program was focused on market validation. The contents and length of each session was the same for each group, but with different logics in the decision-making approaches taught to each group. The first treatment group learnt to develop the BMC using the scientific method; entrepreneurs assigned to the second treatment group learnt to develop the BMC adopting the effectuation method. The control group learnt how to develop the BMC using the same tools of the two treatment groups, but without learning how to use these tools following the scientific method or the effectuation method. We avoided possible contaminations among the three groups following previous best practices used in field experiments regarding decision-making approaches (Camuffo et al. 2024). We observed startups throughout and post-training and we compared differences concerning which side of the BMC entrepreneurs changed and how many times they made pivots to modify the BMC. We considered a left pivot as a change made by blocks that manage infrastructure (key partners, key activities, key structure, cost structure), and a right pivot as a change made by blocks related to value creation (value proposition, customer relationships, customer segments, channels, revenue streams).

3.2 Data collection

We collected data following a two-step approach. The first stage of data collection was performed before the beginning of the pre-acceleration program, during the enrolment phase. Before enrolling in the pre-acceleration program, we asked entrepreneurs to fill out a survey regarding the characteristics of their entrepreneurial idea and about the demography characteristics and human capital of the founding team. This data collection was used to randomize the entrepreneurs between the three treatments and to create part of the independent variables. During the second stage of the data collection, we collected data for 10 points of observation using monthly interviews for ten and a half months. Each interview was led following a semi-structured approach to collect information about the entrepreneurs' decision-making, whether they pivoted or not in terms of changes to their BMC.

3.3 Dependent Variables and Independent Variables

3.3.1 Dependent Variables

As dependent variables, first we used two dependent binary variables to capture which side of the BMC entrepreneurs tend to change. We employed *Left pivot* taking the value 1 whenever an entrepreneur made at least one pivot changing at least one block of the left side of the BMC along the 10 points of observation, 0 otherwise. In a similar vein, we created the variable *Right pivot* taking the value 1 whenever an entrepreneur made at least one pivot changing one block of the right side of the BMC along the 10 points of observation, 0 otherwise.

Second, we investigated the total number of pivots made to the left and right side of the BMC along the 10 points of observation using respectively *Cumulate left pivot*, the cumulate number of pivots made on the left, and *Cumulate right pivot*, the cumulate number of pivots made on the right.

3.3.2 Independent Variables

As independent variable, we employed a categorical variable *Intervention* taking the value 1 whenever an entrepreneur was in the control group, 2 in the scientific group, 3 in the effectuation group.

Finally, we controlled for early termination of the entrepreneurial idea using the control variable *Termination*, taking the value 1 whenever an entrepreneur terminated the entrepreneurial idea, 0 otherwise.

Table 1 presents the descriptive statistics for both the independent and dependent variables.

Table 1**:** Descriptive statistics

Variable	Observations	Mean	Standard deviation	Min	Max	Description
Dependent						
Left pivot	3,388	0.03	0.17	0	1	Binary variable. 1 whenever an entrepreneur makes at least one pivot changing at least one block of the left side of the BMC along the 10 points of observation; 0 otherwise
Right pivot	3,388	0.12	0.33	0	1	Binary variable. 1 whenever an entrepreneur makes at least one pivot changing at least one block of the right side of the BMC along the 10 points of observation; 0 otherwise
Cumulate Left pivot	3,388	0.19	0.50	0	4	Cumulate number of pivots made on the left side of the BMC along the 10 points of observation.
Cumulate Right pivot	3,388	0.91	1.09	0	7	Cumulate number of pivots made on the right side of the BMC along the 10 points of observation.
Independent						
Intervention	308	2.01	0.82	1	3	Binary variable. 1 if an entrepreneur belongs to Control group, 2 to Scientific group, 3 to Effectuation group.
Termination	3,388	0.22	0.41	0	1	Binary variable. 1 whenever an entrepreneur terminates the idea; 0 otherwise.

3.4 Statistical approach

We tested our hypothesis using econometrics analysis, leveraging on data collected along the 10 points of observation. Given the nature of our dependent variables, we first adopted probit model for testing left pivot and right pivot, then negative binomial model for testing the cumulate amount of left pivot and right pivot of early-stage startups. We include a vector of control variables to capture potential differences in the instructors' teaching styles, and we cluster standard errors at both the instructor and intervention levels.

4 Results

Table 2 shows the results of entrepreneurs learning specific entrepreneurial decision-making methods who make pivot decisions on the BMC. Models (1) and (2) present the empirical analysis of Left pivot and Right pivot. Entrepreneurs adopting a scientific method are more likely to make one pivot both on the left and on the right side compared to the control group. On the other hand, entrepreneurs using an effectuation method tend to

make at least one pivot only on the right side compared to the control group. Furthermore, models (3) and (4) consider the cumulate number of pivots. We found that entrepreneurs who learn a scientific method make more pivots on the left side of the BMC along the 10 points of observation. On the other hand, entrepreneurs approaching the effectuation method make more pivots on the right side of the BMC along the 10 points of observation. Collectively, these results confirm our hypothesis showing that different decision-making methods seem to push entrepreneurs to pivot on different parts of their BMC. More precisely, entrepreneurs who learn a scientific method seem to follow a more holistic approach when evolving their entrepreneurial idea, adjusting both the infrastructure side and the value creation side of the BMC. In particular, scientific entrepreneurs tend to make at least one pivot to test their theory with potential customers and, once it is validated, they tend to make more pivots on the left side to focus on infrastructure. On the other hand, entrepreneurs adopting an effectuation method prefer to look for an additional customer segment or identify a different value their idea can offer in response to market conditions rather than modify the initial resources. As a result, both scientific and effectuation entrepreneurs tend to be very careful making strategical pivots showing which side of the BMC they prefer to change.

Table 2: Regressions with *Radical pivot* and *Incremental pivot* as dependent variables.

Variables	**(1)** **DV: Left Pivot** **Probit** **Cross Sections**	**(2)** **DV: Right Pivot** **Probit** **Cross Sections**	**(3)** **DV: Cumulative** **Left Pivot** **Negative Binomial Cross Sections**	**(4)** **DV: Cumulative** **Right Pivot** **Negative Binomial Cross Sections**
Scientific	0.217** (0.023)	0.323*** (0.008)	0.401** (0.017)	0.103 (0.414)
Effectuation	-0.001 (0.994)	0.373* (0.051)	-0.116 (0.633)	0.233** (0.015)
Termination	-0.403** (0.014)	-0.326** (0.023)	-0.883*** (0.000)	-0.614*** (0.000)
Constant	-0.966*** (0.000)	0.193 (0.396)	-1.518*** (0.000)	0.292** (0.015)
Observations	308	308	308	308
Dummies for Instructor	Yes	Yes	Yes	Yes
Clustered errors	Intervention Instructor	Intervention Instructor	Intervention Instructor	Intervention Instructor

Standard errors in parentheses. DV stands for "dependent variable".
***$p < .01$, **$p < .05$, *$p < .10$.

5 Conclusions

This work makes three contributions. First, it contributes to the literature on ET (Alexandria et al., 2014, McKenzie and Woodruff, 2014) and the role of methods applied to entrepreneurial decision-making (Camuffo et al., 2020, Sarasvathy, 2008). In particular, the scientific method and the effectuation method are relevant to create and validate processes. Second, we offer useful insights to literature on pivot decisions in early-stage startups. Pivots are typically associated with organizational flexibility. Thanks to decision-making methods, entrepreneurs learn to consider resources' availability choosing which BMC's blocks are necessary to change, without compromising on flexibility. This consideration has managerial implications that are relevant to practice for entrepreneurs' decisions, concerning resources allocation and long-term development of startups. Third, this study advances entrepreneurial learning process applied to entrepreneurs' pivot decisions involved in RCT experiments. In this way, our findings can support universities and accelerators programs involved in the creation and supply of ETs.

6 Ethics declaration

This research was approved by Bocconi Research Ethics Committee with protocol references SA000105.

7 AI declaration

AI tool ChatGPT was used for proofreading.

References

Agarwal, R., Bacco, F., Camuffo, A., Coali, A., Gambardella, A., Msangi, H., Sonka, S., Temu, A., Waized, B., Wormald, A. (2024) "Does a theory-of-value add value? Evidence from a randomized control trial with Tanzanian entrepreneurs.", *Organization Science*, Vol. 36, Iss. 2.

Alexandria, V., Parton, B., Robb, A. (2014) "Entrepreneurship Education and Training Programs Around the World: Dimensions for Success". *World Bank Publications*.

Bacigalupo, M., Kampylis, P., Punie, Y., Van den Brande, G., and Van den Brande, G., (2016), "EntreComp: The entrepreneurship competence framework", *Publication Office of the European Union*, No. EUR 27939 EN

Bellavitis, C., Tran, M.H. and Wiklund, J. (2025), "Strategic Pivoting in Deep Tech: An Investigation of NSF I-Corps Teams". *Strategic Change*, Vol. 34, pp.305-313.

Camuffo, A., Cordova, A., Gambardella, A., Spina, C. (2020) "A scientific approach to entrepreneurial decision-making: evidence from a randomized control trial", *Strategic Management Journal*, Vol. 66, No. 2, pp. 564-586.

Camuffo, A., Gambardella, A., Messinese, D., Novelli, E., Paolucci, E., Spina, C. (2021) "A Scientific Approach to Innovation Management: Theory and Evidence from Four Field Experiments", *CEPR Discussion Paper*, No. DP15972.

Camuffo, A., Gambardella, A., Messinese, D., Novelli, C., Paolucci, E., Spina, C. (2024) "A scientific approach to entrepreneurial decision-making: Large-scale replication and extension", *Strategic Management Journal*, Vol. 45, Iss. 6, pp. 1209-1237.

Coali, A., Gambardella, A., Novelli, C. (2024) "Scientific decision-making, project selection and longer-term outcomes", *Research Policy*, Vol. 53, Iss. 6.

Coali, A., Novelli, C., Sirigiri, A., Spina, C. (2024) "A scientific approach to decision-making: Key tools and design principles", *Journal of Business Venturing Design*, Vol. 3.

Cohen, S. L., Bingham, C. B., & Hallen, B. L. (2018) The Role of Accelerator Designs in Mitigating Bounded Rationality in New Ventures. *Administrative Science Quarterly*, Vol. 64, No. 4, pp. 810-854.

Felin, T., Gambardella, A., Novelli, E., Zenger, T. (2024) "A scientific method for startups", *Journal of Management*, Vol. 50, No. 8, pp. 3080-3104.

Gonzalez-Tamayo, L. A., Maheshwari, G., Krauss-Delorme, C., Bonomo-Odizzio, A. (2025) "Startup decision-making: integrating trial-and-error with the entrepreneur-as-scientist approach", *Journal of Small Business and Enterprise Development*, Vol. 32, No. 8, pp. 169-196.

Grimes, M. (2018) "The pivot: how founders respond to feedback through idea and identity work", *Academy of Management Journal*, Vol. 61, No. 5, pp.1692–1717.

Kirtley, J., & O'Mahony, S. (2023), "What is a pivot? Explaining when and how entrepreneurial firms decide to make strategic change and pivot", *Strategic Management Journal*, Vol. 44, pp. 197-230.

McDonald, R. M., & Eisenhardt, K. M. (2019) "Parallel Play: Startups, Nascent Markets, and Effective Business-model Design", *Administrative Science Quarterly*, Vol. 65, No. 2, pp. 483-523.

McKenzie, D. and Woodruff, C. (2014) "What Are We Learning from Business Training and Entrepreneurship Evaluations around the Developing World?", *The World Bank Research Observer*, Vol. 29, No. 1, pp. 48–82.

Morris, S., Chad, C., Kistruck, G. M., Lount, R. B., Jr., Thomas, T. E., and Thomas, T. E., 2023, "The impact of growth mindset training on entrepreneurial action among necessity entrepreneurs: Evidence from a randomized control trial", *Strategic Entrepreneurship Journal*, Vol. 17, No. 3, pp. 671–692.

Novelli, E., Spina, C. (2024) "Making business model decisions like scientists: Strategic commitment, uncertainty, and economic performance", *Strategic Management Journal*, Vol. 45, pp. 2642-2695.

Osterwalder, A. (2004) "The Business Model Ontology–A Proposition in A Design Science Approach", Institut d'Informatique et Organisation (Dissertation 173), University of Lausanne, Switzerland.

Ott, T. E. and Hannah, D. P. (2024) "On the Origin of Entrepreneurial Theories: How Entrepreneurs Craft Complex Causal Models with Theorizing and Data", *Strategy Science*, Vol. 9, No. 4, pp. 461-482.

Pillai, S. D., Goldfarb, B., Kirsch, D. A. (2020) "The origins of firm strategy: Learning by economic experimentation and strategic pivots in the early automobile industry", *Strategic Management Journal*, Vol. 41, No. 3, pp. 369–399.

Ries, E. (2011) "How today's entrepreneurs use continuous innovation to create radically successful businesses", The Lean Startup.

Sarasvathy, S. D. (2008) "Effectuation: Elements of Entrepreneurial Expertise", *New Horizons in Entrepreneurship*.

Sarasvathy, S. D. (2024) "Lean Hypotheses and Effectual Commitments: An Integrative Framework Delineating the Methods of Science and Entrepreneurship", *Journal of Management*, Vol. 50, No. 8, pp. 3035-3063.

Slávik, Š. (2019) "The Business Model of Startup—Structure and Consequences", *Administrative Sciences*, Vol. *9*, No. 3, pp. 69

Artificial Intelligence Ambidexterity and Sustainable Business Models: A Configurational Case Study

Francesca Loia[1], Irene Fulco[2], Barbara Aquilani[2], Marcello Martinez[1]
[1] University of Campania Luigi Vanvitelli, Capua, Italy
[2] University of Tuscia, Viterbo, Italy

Abstract

This study explores how artificial intelligence (AI) enables organizational ambidexterity to support sustainability-oriented business models. Building on the knowledge-based view of the firm and organizational ambidexterity theory, it conceptualizes AI as a knowledge-enabling infrastructure that transforms how organizations generate, integrate, and apply knowledge. AI is framed as a dual capability that simultaneously supports knowledge exploitation - through optimization, automation, and efficiency - and knowledge exploration - through experimentation, simulation, and the generation of new insights.

Although prior research on AI and sustainability has expanded, it remains fragmented and often focused on linear applications of AI, overlooking its configurational nature. To address this gap, the study adopts a configurational perspective and investigates how different combinations of AI-enabled exploitation and exploration shape sustainability-oriented business models. A qualitative multiple-case study design is used, based on four multinational firms operating in different industries: Enel, Siemens, Unilever, and Nestlé. Data were collected from secondary sources, including sustainability reports, annual reports, and corporate strategy documents, and analyzed through within-case and cross-case comparison.

The findings identify four distinct archetypes of AI-enabled knowledge ambidexterity: (1) an ambidextrous sustainability transformer characterized by high levels of both exploitation and exploration; (2) an efficiency-oriented sustainability optimizer dominated by exploitation; (3) an innovation-oriented sustainability explorer driven mainly by exploration; and (4) a limited AI adoption configuration with weak integration of both processes. Each configuration corresponds to a different sustainability-oriented business model logic, ranging from system-level transformation to incremental efficiency gains and exploratory but less scalable innovation.

The results show that sustainability outcomes are not determined by AI adoption alone, but by how AI-enabled knowledge processes are configured and aligned with organizational structures and business model logic. The study highlights equifinality, demonstrating that different configurations can lead to sustainability through distinct pathways.

Theoretically, the paper contributes by reframing AI-enabled ambidexterity as a configurational knowledge capability and integrating ambidexterity theory with sustainable business model research. Practically, it suggests that managers should move beyond fragmented AI initiatives and instead design integrated systems that balance exploitation and exploration within coherent organizational and strategic architectures.

Keywords: Artificial intelligence; ambidexterity; business model; sustainable business model

Paper type: Academic Research Paper

1 Introduction

The increasing urgency of environmental, social, and governance (ESG) challenges is fundamentally reshaping how organizations create, manage, and leverage knowledge to support sustainable value creation. In this context, Artificial Intelligence (AI) is emerging as a critical knowledge-enabling technology, transforming how organizations generate, integrate, and exploit knowledge for sustainability-oriented strategies (George et al., 2021).

Despite growing scholarly attention, existing research on AI and sustainability remains fragmented. Prior studies tend to adopt linear perspectives, focusing either on efficiency-oriented applications (e.g., optimization, cost reduction) or on innovation-driven initiatives (Bai et al., 2020). Such approaches overlook the inherently ambidextrous nature of AI when understood as a knowledge-based organizational capability.

Drawing on the knowledge-based view of the firm (Grant, 1996) and organizational ambidexterity theory (March, 1991), this study argues that AI enables organizations to simultaneously engage in knowledge exploitation and exploration. While exploitation supports sustainability through efficiency, optimization, and reuse of existing knowledge, exploration enables experimentation, learning, and the generation of novel sustainability solutions. Importantly, these processes are not independent but interdependent and configurational in nature (Raisch et al., 2009).

This configurational perspective is particularly relevant in sustainability contexts, where organizations must address complex, systemic, and long-term challenges requiring both incremental improvements and

transformative innovation. However, there remains limited understanding of how different configurations of AI-enabled knowledge processes shape sustainability-oriented business models.

To address this gap, this study investigates the following research question:

How do different configurations of AI-enabled exploitation and exploration shape sustainability-oriented business models?

Adopting a configurational case study approach based on four multinational firms - Enel, Siemens, Unilever, and Nestlé - this study identifies archetypal patterns of AI-enabled knowledge ambidexterity and links them to distinct sustainable business model logics.

This study contributes to the literature in three ways. First, it conceptualizes AI ambidexterity as a knowledge-based and configurational capability rather than a linear technological investment. Second, it integrates ambidexterity theory with sustainable business model research by highlighting the role of business model alignment. Third, it demonstrates equifinality in AI-enabled sustainability, showing how different configurations of knowledge processes can lead to distinct sustainability pathways.

2 Theoretical Background

2.1 AI as a knowledge-enabling infrastructure

AI is increasingly conceptualized not merely as a technological tool, but as a knowledge-enabling infrastructure that fundamentally reshapes how organizations generate, process, and leverage knowledge (Ferraris et al., 2019). Unlike traditional information systems, AI systems are capable of autonomously identifying patterns, generating predictions, and supporting decision-making processes, thereby transforming raw data into actionable organizational knowledge (Barney, 2021; George et al., 2021).

From a knowledge-based perspective, the firm can be understood as an entity specialized in the integration and application of knowledge (Nonaka, 1994; Grant, 1996). Within this view, AI extends the firm's knowledge-processing capabilities by enhancing key processes such as knowledge codification, integration, and recombination (Alavi & Leidner, 2001; Davenport & Ronanki, 2018). In particular, AI enables the transformation of dispersed and unstructured data into structured knowledge, facilitating its reuse across organizational domains (Brynjolfsson & McElheran, 2016; Jarrahi, 2018).

This role becomes especially critical in sustainability contexts, where organizations must process complex, multidimensional information related to environmental impact, social responsibility, and governance practices. AI allows firms to handle such complexity by integrating heterogeneous data sources and generating insights that support sustainability-oriented decision-making (Vinuesa et al., 2020; Urbanovič & Holubčík, 2026).

However, the impact of AI on organizational knowledge processes is not unidimensional. While AI can enhance efficiency through optimization and automation, it also enables experimentation and the generation of novel knowledge through simulations, predictive modeling, and scenario analysis (Ransbotham et al., 2017; Raisch & Krakowski, 2021). This dual role suggests that AI should be understood as a technology that simultaneously supports both the exploitation and exploration of knowledge.

2.2 Organizational ambidexterity as knowledge process integration

The distinction between exploitation and exploration represents a foundational lens for understanding organizational learning and adaptation (March, 1991). Exploitation refers to activities aimed at refining and extending existing knowledge, emphasizing efficiency, reliability, and incremental improvement. Exploration, in contrast, involves experimentation, variation, and the pursuit of new knowledge, often associated with uncertainty and risk.

Traditional perspectives have framed exploitation and exploration as competing activities, requiring organizations to balance trade-offs between efficiency and innovation (Levinthal & March, 1993; He & Wong, 2004). However, more recent research has challenged this dichotomy, emphasizing the interdependence and configurational nature of these processes (Raisch et al., 2009; O'Reilly & Tushman, 2013; Zhao & Gao, 2024). Rather than being mutually exclusive, exploitation and exploration can be understood as complementary and mutually reinforcing components of organizational learning (Liu, 2024).

From a knowledge-based perspective, ambidexterity reflects the organization's ability to manage and recombine different types of knowledge processes (Andriopoulos & Lewis, 2009; Smith & Lewis, 2011). Exploitation supports the accumulation and refinement of existing knowledge, while exploration enables the generation of new knowledge and the discovery of novel opportunities. The effectiveness of ambidexterity thus

depends not only on the presence of both processes, but on how they are combined and embedded within organizational structures and routines.

AI technologies play a critical role in enabling this integration. AI-driven exploitation enhances knowledge efficiency through automation, predictive analytics, and optimization (McAfee & Brynjolfsson, 2017; Davenport et al., 2020). Conversely, AI-driven exploration facilitates knowledge creation by enabling simulation, experimentation, and the exploration of alternative scenarios (Cockburn et al., 2018; Haefner et al., 2021). Rather than operating as independent processes, exploitation and exploration are recursively intertwined, as efficiency-oriented learning processes enable the conditions under which novel variation and innovation can emerge.

2.3 Sustainable business models and knowledge alignment

Sustainable business models extend traditional business model frameworks by integrating economic, environmental, and social value creation (Boons & Lüdeke-Freund, 2013; Bocken et al., 2014; Lüdeke-Freund & Dembek, 2017). These models require organizations to rethink how value is created, delivered, and captured, incorporating sustainability considerations into core strategic and operational processes (Fulco, 2025).

Achieving sustainability outcomes involves more than adopting isolated technologies or practices. Instead, it requires the alignment of organizational capabilities, knowledge processes, and strategic objectives (Stubbs & Cocklin, 2008; Schaltegger et al., 2016). In particular, sustainability challenges often involve systemic complexity, long time horizons, and multiple stakeholders, requiring organizations to integrate diverse forms of knowledge and continuously adapt their strategies.

In this context, knowledge plays a central role in enabling sustainable business models (Adams et al., 2016; Raina, 2026). Organizations must be able to leverage existing knowledge to improve efficiency and reduce environmental impact, while also generating new knowledge to develop innovative solutions and business practices. This dual requirement highlights the importance of ambidextrous knowledge processes.

However, existing research has largely overlooked how these knowledge processes are configured within organizations and how they interact with business model structures. As a result, there remains limited understanding of how AI-enabled knowledge processes contribute to sustainability outcomes through their alignment with business model logic.

2.4 A Configurational view of AI-enabled knowledge ambidexterity

Building on the previous sections, this study proposes a configurational perspective on AI-enabled knowledge ambidexterity. Rather than treating exploitation and exploration as independent or additive dimensions, this perspective emphasizes their interdependence and the importance of their alignment within organizational contexts.

AI-enabled exploitation and exploration represent distinct but interconnected knowledge processes. Their interaction shapes the organization’s ability to generate, integrate, and apply knowledge for sustainability (Raisch & Birkinshaw, 2008; O’Reilly & Tushman, 2013). Importantly, different combinations of these processes may lead to different organizational outcomes, suggesting the presence of multiple pathways to sustainability.

This perspective aligns with configurational approaches, which emphasize that organizational outcomes result from specific combinations of interdependent elements rather than from individual factors considered in isolation (Ragin, 2009; Fiss, 2011). In this view, sustainability-oriented business models emerge from coherent configurations of AI-enabled knowledge processes embedded within the organization.

Accordingly, this study examines how different configurations of AI-enabled exploitation and exploration give rise to distinct sustainability-oriented business model archetypes.

3 Methodology

3.1 Research design

This study adopts a qualitative, configurational multiple-case study design to investigate how different combinations of AI-enabled knowledge processes shape sustainability-oriented business models. A configurational approach is particularly suitable for this research, as it enables the analysis of complex causal patterns characterized by interdependence, equifinality, and non-linearity (Ragin, 2009; Fiss, 2011). While configurational theory is often associated with set-theoretic methods such as fuzzy-set qualitative comparative analysis (fsQCA) (Ragin, 2009), this study adopts a qualitative configurational approach aimed at theory building and pattern identification rather than formal set-theoretic analysis.

Rather than focusing on the net effects of individual variables, the study examines how specific configurations of AI-enabled exploitation and exploration jointly produce distinct sustainability-oriented outcomes. This perspective is particularly relevant given the systemic and multi-dimensional nature of sustainability challenges, which require the alignment of technological capabilities, knowledge processes, and business model structures.

The multiple-case study design enhances analytical robustness by enabling systematic cross-case comparison and pattern identification (Yin, 2018; Eisenhardt & Graebner, 2007). In line with configurational theorizing, the objective is not statistical generalization, but the development of analytical generalizations in the form of empirically grounded configurational archetypes, synthesized in a comparative table.

3.2 Case selection

The empirical analysis is based on four multinational firms: Enel, Siemens, Unilever, and Nestlé. These cases were selected through purposeful theoretical sampling to ensure variation across industries while maintaining relevance to the research objective. The selection was guided by three criteria:
(1) the presence of significant or emerging AI adoption in organizational processes,
(2) a clearly articulated strategic commitment to sustainability objectives, and
(3) the availability of rich and reliable secondary data sources.

The selected firms operate in distinct sectors - energy, industrial manufacturing, consumer goods, and food and beverage - allowing for cross-industry comparison. This heterogeneity increases the likelihood of identifying diverse configurations of AI-enabled knowledge processes and strengthens the analytical generalizability of the findings (Eisenhardt, 1989).

3.3 Data collection

Data were collected from multiple secondary sources to ensure richness, depth, and triangulation. These sources include sustainability reports, integrated annual reports, corporate strategy documents, AI and digital transformation reports, investor presentations, and publicly available executive communications.

Secondary data are particularly appropriate for this study, as multinational firms systematically disclose both sustainability and digital transformation initiatives, providing detailed insights into organizational practices and strategic orientations.

Data collection focused specifically on identifying AI-enabled activities related to organizational knowledge processes, as well as their integration into sustainability strategies and business model configurations.

Triangulation across multiple sources enhances construct validity and reduces interpretive bias (Yin, 2018; Denzin, 1978, 2017). Furthermore, the analysis was conducted by two researchers, who independently examined the data and subsequently compared and discussed their interpretations, thereby strengthening the reliability and robustness of the findings.

3.4 Data analysis

The data analysis followed a structured, iterative, and comparative process consistent with configurational research logic.

First, AI-enabled activities were identified and coded inductively from the dataset, focusing on their role in organizational knowledge processes such as data processing, optimization, decision-making, experimentation, and learning.

Second, these activities were categorized as either AI-enabled exploitation or AI-enabled exploration, following established definitions in the ambidexterity literature (March, 1991).

Third, within-case analysis was conducted to assess the relative dominance and interaction of exploitation and exploration within each organization, as well as their embedding in sustainability-oriented business model structures. This step enabled the construction of detailed case profiles.

Fourth, a cross-case comparison was performed to identify similarities and differences across cases. Through iterative comparison and abstraction, cases were grouped into higher-order configurational archetypes.

Finally, these archetypes were synthesized in a comparative table (Table 1), which summarizes the identified configurations, their dominant knowledge processes, and associated sustainability-oriented business model logics.

Importantly, ambidexterity is conceptualized not as a continuous or additive construct, but as a configurational property emerging from the interaction and alignment of AI-enabled knowledge processes

within organizational contexts. This enables a more nuanced understanding of how sustainability outcomes are generated through interdependent mechanisms rather than isolated factors.

4 Findings

The analysis reveals four distinct configurations of AI-enabled knowledge ambidexterity, each corresponding to a specific sustainability-oriented business model logic. Table 1 presents these configurational archetypes, illustrating how AI-enabled exploitation and exploration interact to shape underlying knowledge processes and their alignment with sustainability-oriented business model logics.

Table1. Configurational archetypes of AI-enabled knowledge ambidexterity

Archetype	Firm	AI exploitation	AI exploration	Dominant knowledge processes	Business model logic	Sustainability outcome
Ambidextrous sustainability transformer	Enel	High	High	Knowledge integration, recombination, continuous learning	Transformative, system-level	Balanced efficiency and innovation; long-term sustainability transition
Efficiency-oriented sustainability optimizer	Siemens	High	Low	Knowledge codification, optimization, reuse	Efficiency-driven	Incremental sustainability improvements; operational excellence
Innovation-oriented sustainability explorer	Unilever	Low-Moderate	High	Knowledge creation, experimentation, learning	Innovation-driven	Exploratory sustainability initiatives; limited scalability
Limited AI adoption	Nestlé	Low	Low	Fragmented knowledge processes	Conventional/ transitional	Limited sustainability impact; low transformation

Source: Authors' elaboration

The first configuration, labelled Ambidextrous sustainability transformer and exemplified by Enel, is characterized by high levels of both exploitation and exploration. In this case, AI is deployed not only to optimize existing energy infrastructures through predictive maintenance, smart grid management, and real-time monitoring of distributed assets, but also to support exploratory capabilities such as digital twin infrastructures and advanced scenario simulations for the integration of renewable energy sources into the grid. This combination enables Enel to simultaneously pursue operational efficiency and systemic transformation, particularly in the context of energy transition and decarbonization strategies. The dual orientation supports the development of a transformative business model that balances short-term efficiency gains with long-term infrastructural innovation. The tight integration of exploitation and exploration fosters continuous feedback loops between operational data and strategic foresight, positioning sustainability as a core strategic driver embedded within the organization's socio-technical architecture.

A second configuration, the Efficiency-oriented sustainability optimizer, represented by Siemens, reflects a strong emphasis on AI-enabled exploitation combined with limited exploration. Siemens primarily leverages AI in industrial and engineering contexts to enhance operational efficiency, optimize manufacturing systems, and improve resource utilization across complex industrial environments. Applications such as predictive maintenance for industrial equipment, process automation, and energy efficiency optimization dominate the AI portfolio, while exploratory initiatives remain more selective and often confined to specific innovation units. As a result, sustainability outcomes are largely driven by incremental efficiency improvements and the systematic refinement of existing engineering and process knowledge. The corresponding business model is strongly anchored in industrial reliability, performance optimization, and scalable efficiency, with relatively limited engagement in experimental or radically innovative sustainability pathways.

In contrast, the Innovation-oriented sustainability explorer, exemplified by Unilever, displays a configuration in which exploration dominates over exploitation. Here, AI is primarily used to generate new knowledge related to sustainable sourcing strategies, product innovation, and evolving consumer behavior patterns. Machine learning and advanced analytics are applied to identify emerging sustainability trends, support portfolio redesign, and guide experimentation in areas such as circular economy initiatives and responsible consumption models. This orientation supports a business model centered on experimentation, rapid learning, and portfolio

diversification in sustainability-oriented offerings. However, the relatively lower level of exploitation limits the scalability of these initiatives, as many remain at pilot or project level, with partial integration into standardized operational routines and global supply chain processes.

Finally, the Limited AI adoption configuration, represented by Nestlé, is characterized by low levels of both exploitation and exploration. In this case, AI initiatives appear more fragmented and less systematically embedded within core organizational knowledge processes. While some applications of digital analytics and data-driven decision support exist - particularly in areas such as supply chain monitoring or demand forecasting - these initiatives do not yet constitute a fully integrated AI-enabled knowledge system. Consequently, sustainability efforts remain largely grounded in established managerial practices and incremental improvements rather than in AI-driven transformation. This configuration highlights that the absence of a coherent integration between AI-enabled knowledge processes and business model structures significantly constrains the potential for sustainability-oriented innovation and systemic impact.

5 Discussion

The findings offer a configurational explanation of how AI-enabled knowledge processes shape sustainability-oriented business models, moving beyond linear and technology-centric interpretations of AI impact. Rather than viewing artificial intelligence as an exogenous driver of sustainability performance, the analysis shows that sustainability emerges from the systemic alignment of AI-enabled knowledge processes within organizational architectures and business model logics.

A first theoretical contribution lies in the reconceptualization of AI-enabled ambidexterity as a configurational and relational knowledge capability. Consistent with knowledge-based and ambidexterity perspectives (March, 1991; Grant, 1996; O'Reilly & Tushman, 2013), the findings show that exploitation and exploration are not independent organizational choices, but mutually constitutive processes embedded in AI-mediated knowledge systems.

AI plays a dual role in this configuration: it simultaneously enhances the efficiency of knowledge exploitation (e.g., optimization, prediction, automation) and expands the scope of knowledge exploration (e.g., simulation, pattern discovery, scenario building). However, the study demonstrates that value does not stem from AI in isolation, but from how these two logics are configured, balanced, and institutionally embedded within organizational practices.

This shifts the theoretical focus from AI as a capability to AI as a meta-enabler of organizational knowledge configurations, where sustainability outcomes emerge from interaction effects rather than additive contributions.

A second contribution concerns the identification of equifinality in AI-enabled sustainability transformation. The results show that multiple distinct configurations can lead to sustainability-oriented outcomes, confirming configurational logic (Ragin, 2009; Fiss, 2011). However, this equifinality is not homogeneous.

Instead, different configurations produce qualitatively distinct sustainability pathways, ranging from efficiency-driven incrementalism to exploratory innovation and system-level transformation. This suggests that sustainability is not a single outcome, but a plural and path-dependent phenomenon shaped by underlying knowledge architectures.

Accordingly, the study extends prior research on sustainable business models (Boons & Lüdeke-Freund, 2013; Bocken et al., 2014) by showing that heterogeneity is a structural feature of how AI-enabled firms pursue sustainability.

A third key insight is the role of business model alignment as a critical mediating mechanism linking AI-enabled knowledge processes to sustainability outcomes. The findings suggest that AI-driven exploitation and exploration generate strategic value only when they are embedded within coherent configurations of value creation, delivery, and capture.

This extends the sustainable business model literature by shifting attention from "what firms do" to how firms structurally integrate AI-enabled knowledge processes into their business logic. In other words, AI does not automatically translate into sustainability benefits; rather, it requires architectural alignment between knowledge processes and business model design.

Without such alignment, AI remains fragmented, underutilized, or confined to localized efficiency gains, limiting its systemic sustainability impact.

Finally, the findings highlight a structural paradox of imbalance in AI-enabled ambidexterity. Configurations dominated by exploitation tend to generate efficiency gains but risk reinforcing incrementalism and path dependency. Conversely, configurations dominated by exploration foster innovation and learning but often suffer from weak operationalization and limited scalability.

This tension reflects broader ambidexterity paradoxes (Raisch & Birkinshaw, 2008; Smith & Lewis, 2011), but the study extends this logic by showing that AI intensifies both sides of the tension. AI amplifies exploitation through automation and optimization, while simultaneously lowering the cost of exploration through simulation and data-driven experimentation.

As a result, AI does not resolve the ambidexterity paradox - it reconfigures and potentially amplifies it, making organizational alignment even more critical for achieving sustainable outcomes.

6 Conclusions

This study advances the literature on artificial intelligence and sustainability by conceptualizing AI-enabled ambidexterity as a configurational knowledge capability. By integrating insights from the knowledge-based view, organizational ambidexterity theory, and sustainable business model research, it explains how organizations leverage AI not as a standalone technology, but as an enabler of complex knowledge configurations for sustainable value creation.

The findings demonstrate that the effectiveness of AI in sustainability contexts does not depend on technological adoption per se, but on how AI is embedded within organizational knowledge processes and aligned with business model structures. In this sense, sustainability emerges from the coherence between technological affordances, knowledge dynamics, and business model design, rather than from isolated digital initiatives. This shifts the analytical focus from AI as a tool to AI as a structural enabler of organizational configuration.

From a managerial perspective, the study suggests that organizations should move beyond fragmented AI implementations and instead deliberately orchestrate AI-enabled exploitation and exploration as interconnected processes. The results highlight the importance of embedding these processes within organizational routines, decision-making systems, and strategic architectures. In particular, managers should be aware that imbalances in ambidexterity may generate unintended consequences: overemphasis on exploitation can lead to incrementalism and rigidity, while excessive exploration may hinder scalability and operational integration.

Despite these contributions, the study has limitations that open avenues for future research. First, the analysis is based on a limited number of cases, which, while appropriate for configurational theory-building, constrains broader generalizability. Second, the reliance on secondary data may not fully capture micro-level organizational dynamics, informal practices, or internal decision-making processes underlying AI adoption and use. Future research could address these limitations by adopting longitudinal designs to capture the temporal dynamics of AI-enabled ambidexterity, as well as by incorporating primary data such as interviews and ethnographic observations.

Further research could also extend the configurational framework proposed in this study by testing it across a larger and more diverse set of organizations, industries, and institutional contexts. Importantly, the equifinality emerging from our findings - namely, the existence of multiple, functionally equivalent configurations leading to sustainability outcomes - calls for more systematic empirical validation. Future studies could employ methodologies such as fsQCA to rigorously test alternative configurational pathways and assess their relative consistency and coverage across contexts. In addition, comparative studies could explore how national or regulatory environments shape the emergence of different AI-enabled sustainability configurations. Finally, future work could investigate the micro-foundations of AI-enabled ambidexterity, particularly the role of managerial cognition, organizational culture, and data governance in shaping configurational outcomes.

7 Ethics Declaration

Ethics approval was not required for this study.

8 AI Declaration

No AI tools were used in the development of this paper.

References

Adams, R., Jeanrenaud, S., Bessant, J., Denyer, D., & Overy, P. (2016). Sustainability-oriented innovation: A systematic review. *International Journal of Management Reviews*, *18*(2), 180-205.

Alavi, M., & Leidner, D. E. (2001). Knowledge management and knowledge management systems: Conceptual foundations and research issues1, 2. *MIS quarterly*, *25*(1), 107-136.

Andriopoulos, C., & Lewis, M. W. (2009). Exploitation-exploration tensions and organizational ambidexterity: Managing paradoxes of innovation. *Organization science*, *20*(4), 696-717.

Bai, C., Dallasega, P., Orzes, G., & Sarkis, J. (2020). Industry 4.0 technologies assessment: A sustainability perspective. *International Journal of Production Economics*, *229*, 107776.

Barney, J. B. (2021). The emergence of resource-based theory: A personal journey. *Journal of Management*, *47*(7), 1663-1676.

Bocken, N. M. P., Short, S. W., Rana, P., & Evans, S. (2014). A literature and practice review to develop sustainable business model archetypes. *Journal of Cleaner Production*, *65*, 42–56.

Boons, F., & Lüdeke-Freund, F. (2013). Business models for sustainable innovation: State-of-the-art and steps towards a research agenda. *Journal of Cleaner Production*, *45*, 9–19.

Brynjolfsson, E., & McElheran, K. (2016). The rapid adoption of data-driven decision-making. *American Economic Review*, *106*(5), 133-139.

Cockburn, I. M., Henderson, R., & Stern, S. (2018). The impact of artificial intelligence on innovation: An exploratory analysis. In *The economics of artificial intelligence: An agenda* (pp. 115-146). University of Chicago Press.

Davenport, T. H., & Ronanki, R. (2018). Artificial intelligence for the real world. *Harvard business review*, *96*(1), 108-116.

Davenport, T., Guha, A., Grewal, D., & Bressgott, T. (2020). How artificial intelligence will change the future of marketing. *Journal of the academy of marketing science*, *48*(1), 24-42.

Denzin, N. K. (1978). Triangulation: A case for methodological evaluation and combination. *Sociological methods*, *56*, 339-357.

Denzin, N. K. (2017). *The research act: A theoretical introduction to sociological methods*. Routledge.

Eisenhardt, K. M. (1989). Building theories from case study research. *Academy of management review*, *14*(4), 532-550.

Eisenhardt, K. M., & Graebner, M. E. (2007). Theory building from cases: Opportunities and challenges. *Academy of management journal*, *50*(1), 25-32.

Ferraris, A., Mazzoleni, A., Devalle, A., & Couturier, J. (2019). Big data analytics capabilities and knowledge management: Impact on firm performance. *Management Decision*, *57*(8), 1923–1936.

Fiss, P. C. (2011). Building better causal theories: A fuzzy set approach to typologies in organization research. *Academy of management journal*, *54*(2), 393-420.

Fulco, I. (2025). *Business model innovation e co-creazione di valore nelle imprese: il ruolo delle piattaforme digitali*. [Assago]: CEDAM.

George, G., Merrill, R. K., & Schillebeeckx, S. J. D. (2021). Digital sustainability and entrepreneurship: How digital innovations are helping tackle climate change and sustainable development. *Entrepreneurship Theory and Practice*, *45*(5), 999–1027.

Grant, R. M. (1996). Toward a knowledge-based theory of the firm. *Strategic Management Journal*, *17*(S2), 109–122.

Haefner, N., Wincent, J., Parida, V., & Gassmann, O. (2021). Artificial intelligence and innovation management: A review, framework, and research agenda☆. *Technological forecasting and social change*, *162*, 120392.

He, Z. L., & Wong, P. K. (2004). Exploration vs. exploitation: An empirical test of the ambidexterity hypothesis. *Organization Science*, *15*(4), 481–494.

Jarrahi, M. H. (2018). Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making. *Business horizons*, *61*(4), 577-586.

Levinthal, D. A., & March, J. G. (1993). The myopia of learning. *Strategic management journal*, *14*(S2), 95-112.

Liu, S. (2024). Algorithmic ambidexterity: Rethinking exploration and exploitation in the age of AI. *Available at SSRN 5291744*.

Lüdeke-Freund, F., & Dembek, K. (2017). Sustainable business model research and practice: Emerging field or passing fancy? *Journal of Cleaner Production*, *168*, 1668–1678.

March, J. G. (1991). Exploration and exploitation in organizational learning. *Organization Science*, *2*(1), 71–87.

McAfee, A., & Brynjolfsson, E. (2017). *Machine, platform, crowd: Harnessing our digital future*. WW Norton & Company.

Nonaka, I. (1994). A dynamic theory of organizational knowledge creation. *Organization science*, *5*(1), 14-37.

O'Reilly, C. A., & Tushman, M. L. (2013). Organizational ambidexterity: Past, present, and future. *Academy of Management Perspectives*, *27*(4), 324–338.

Ragin, C. C. (2009). *Redesigning social inquiry: Fuzzy sets and beyond*. University of Chicago press.

Raina, K., Sharma, G. D., Taheri, B., Dev, D., & Chavriya, S. (2026). Artificial intelligence-driven management: Bridging innovation, knowledge creation, and sustainable business practices. *Journal of Innovation & Knowledge*, *11*, 100860.

Raisch, S., & Birkinshaw, J. (2008). Organizational ambidexterity: Antecedents, outcomes, and moderators. *Journal of management*, *34*(3), 375-409.

Raisch, S., & Krakowski, S. (2021). Artificial intelligence and management: The automation–augmentation paradox. *Academy of management review*, *46*(1), 192-210.

Raisch, S., Birkinshaw, J., Probst, G., & Tushman, M. L. (2009). Organizational ambidexterity: Balancing exploitation and exploration for sustained performance. *Organization Science*, *20*(4), 685–695.

Ransbotham, S., Kiron, D., Gerbert, P., & Reeves, M. (2017). Reshaping business with artificial intelligence: Closing the gap between ambition and action. *MIT sloan management review*, *59*(1).

Schaltegger, S., Hansen, E. G., & Lüdeke-Freund, F. (2016). Business models for sustainability: Origins, present research, and future avenues. *Organization & environment*, *29*(1), 3-10.

Smith, W. K., & Lewis, M. W. (2011). Toward a theory of paradox: A dynamic equilibrium model of organizing. *Academy of management Review*, *36*(2), 381-403.

Stubbs, W., & Cocklin, C. (2008). Conceptualizing a “sustainability business model”. *Organization & environment*, *21*(2), 103-127.

Urbanovič, M., & Holubčík, M. (2026). Artificial Intelligence in Managerial Decision-Making for Sustainable Business Models: A Systematic Literature Review. *Systems*, *14*(3), 245.

Vinuesa, R., Azizpour, H., Leite, I., Balaam, M., Dignum, V., Domisch, S., ... & Fuso Nerini, F. (2020). The role of artificial intelligence in achieving the Sustainable Development Goals. *Nature communications*, *11*(1), 233.

Yin, R. K. (2018). *Case study research and applications* (Vol. 6). Thousand Oaks, CA: Sage.

Zhao, J., & Gao, P. (2024). Configurations of ambidextrous innovation and its performance implication in the context of digital transformation. *Systems*, *12*(2), 60.

Hybrid-Knowledge Management for Sustainable Competitiveness: AI-Human Synergy as Value Enabler

Andrea Gargiulo, Daniele Leone, Salvatore Scognamiglio, Petrella Raffaele Clemente
University of Naples "Parthenope", Naples, Italy

Abstract

This paper examines how firms can govern hybrid-knowledge management to transform sustainability information into customer-perceived competitive value. The study starts from a practical problem related to sustainability investments, which often fail to generate advantage. This occurs because customers receive fragmented or poorly translated signals. In response, the paper defines hybrid-knowledge management as the deliberate design and governance of customer-facing hybrid sustainability knowledge flows. These exchanges combine AI-generated cues with human-mediated interpretation provided by influencers, frontline employees, and brand communicators. The research asks how sustainable competitiveness can be strengthened through hybrid-knowledge mechanisms involving both humans and AI agents. A qualitative multi-case study design is adopted, involving four cases selected across the food and fashion industries. This setting was chosen because both sectors are highly exposed to sustainability-sensitive consumer choice and feature visible customer journeys in which knowledge is mediated through both digital and human channels. Data collection combined direct observation of customer-facing interfaces and campaign touchpoints with documentary secondary sources, enabling triangulation across firms, channels, and touchpoints. Data were analysed inductively through the Gioia methodology. The findings show that sustainable competitiveness is shaped by three interrelated levers. First, claim integrity, which rests on sustainability cues being auditable, standardised, and accompanied by human explanation. Second, choice guidance, which requires sustainability information to be translated into decision-ready comparisons at the point of choice. Third, relational sensemaking, which depends on customers' ability to resolve ambiguity through trusted human interaction and receive continuity after purchase. Across the cases, no instance of full alignment was observed. Two-lever configurations generated partial clarity, while one-lever configurations produced sustainability blur, scepticism, or weak conversion. The paper contributes by theorising hybrid-knowledge management as a strategic capability for orchestrating customer perceptions of sustainability, rather than as a purely internal knowledge process. It also offers actionable implications for managers, specifically to align verification, explainability, and human interpretation across touchpoints, and for policymakers seeking stronger traceability and comparability standards in both digital and human communication channels.

Keywords: Hybrid-knowledge management; artificial intelligence; sustainable competitiveness; customer perception

Paper type: Academic Research Paper

1 Introduction

Sustainable competitiveness increasingly depends on the way firms mobilise knowledge assets to shape market understanding, customer judgement, and organisational adaptation over time (Grant, 1996; Nonaka and Toyama, 2003). Recent advances in artificial intelligence have amplified this challenge by extending the scale, speed, and standardisation with which firms can create, store, retrieve, and apply knowledge in customer-facing settings (Jarrahi et al., 2023). At the same time, the business returns from responsible conduct remain contingent on customer awareness, because sustainability investments do not create value when customers neither notice nor understand them (Servaes and Tamayo, 2013; Du, Bhattacharya and Sen, 2010). This communication problem becomes more severe when stakeholders suspect exaggeration, symbolic signalling, or selective disclosure, since sustainability claims that lack credible support can trigger greenwashing inferences rather than competitive gains (Delmas and Burbano, 2011). From a strategic standpoint, sustainable offerings only become sources of advantage when firms connect responsible practices with clear value propositions that customers can interpret and use in choice situations (Kahupi et al., 2021; Hermundsdottir and Aspelund, 2022). In practice, sustainability-related knowledge now reaches customers through two intertwined mediation paths: human actors such as influencers and frontline employees, and AI-enabled interfaces such as recommender systems, labels, filters, and decision-support cues (Lou and Yuan, 2019; Cambra-Fierro, Melero-Polo and Vázquez-Carrasco, 2014). These AI-enabled cues may raise perceived objectivity and decision relevance, but they also reshape human judgement in ways that depend on trust, interpretability, and the context of decision support (Logg, Minson and Moore, 2019; Poszler and Lange, 2024). Despite growing work on knowledge management,

algorithmic advice, customer awareness, and sustainability strategy, research still says little about how firms should intentionally combine AI-mediated and human-mediated knowledge as a strategic capability for sustainable competitiveness. In this paper, hybrid-knowledge management is therefore defined as the deliberate design and governance of customer-facing sustainability knowledge flows that combine AI-generated cues with human-mediated interpretation across touchpoints in the customer journey. This definition addresses two limitations highlighted in current managerial debates. First, it specifies what is hybrid in operational terms. Second, it links the topic more directly to strategy by treating the orchestration of knowledge flows as a deliberate managerial capability rather than a background communication condition. The study accordingly investigates the following research question: How can sustainable competitiveness be increased through hybrid-knowledge mechanisms involving humans and AI agents?

2 Literature review

2.1 Hybrid-knowledge management and customer-facing knowledge flows

Knowledge-based theory explains competitive advantage in terms of how firms integrate and deploy knowledge assets more effectively than rivals (Grant, 1996). Knowledge creation theory further shows that value emerges when distributed experience is synthesised into actionable meaning rather than merely accumulated as information (Nonaka and Toyama, 2003). As contemporary work extends these insights to partnerships between AI systems and humans combining codification with interpretation, the concept of hybrid-knowledge management emerges (Lu, Nerode and Subrahmanian, 1996; Jarrahi et al., 2023). In market-facing contexts, frontline employees mediate knowledge by helping customers interpret offers, evaluate trade-offs, and make sense of brand claims in interaction (Cambra-Fierro, Melero-Polo and Vázquez-Carrasco, 2014). Influencer communication similarly depends on perceived message value and source credibility, which means that human-mediated sustainability claims are shaped by trust as much as by information content (Lou and Yuan, 2019). Follower cues and perceived popularity can strengthen influencer effectiveness, especially in image-driven categories such as fashion where social proof often substitutes for technical explanation (De Veirman, Cauberghe and Hudders, 2017).

2.2 AI-mediated knowledge and sustainability perception

AI-mediated knowledge operates through scalable recommendations, rankings, and decision-support outputs that can increase reliance on algorithmic judgement when users perceive advice as useful and objective (Logg, Minson and Moore, 2019). However, resistance emerges when users observe errors and conclude that algorithmic systems are less forgivable or less adaptive than human advisors (Dietvorst, Simmons and Massey, 2015). One design response is to demonstrate that algorithmic systems learn over time, because visible learning can offset the negative effects of familiarity on reliance (Berger et al., 2021). Acceptance also depends on system transparency, since explainability indirectly shapes trust and thereby affects whether intelligent systems are used in consequential tasks (Wanner et al., 2022). Research on intelligent decision-support systems similarly shows that ethical decision support requires intelligible rationales rather than opaque optimisation alone (Poszler and Lange, 2024).

2.3 Sustainability communication, competitiveness and the analytical framework

Corporate Social Responsibility (CSR) efforts create more value when customers are aware of them, because stakeholder response mediates the effect of responsible conduct on firm outcomes (Servaes and Tamayo, 2013). CSR communication research therefore emphasises that awareness must be paired with favourable attributions if firms are to capture reputational and behavioural returns from responsible action (Du, Bhattacharya and Sen, 2010). When communication outruns substantiation, greenwashing becomes more likely and can erode confidence in green products, green firms, and sustainability claims more broadly (Delmas and Burbano, 2011). Consumer scepticism toward sustainability claims is heightened by previous greenwashing incidents and can even make quieter communication strategies appear more credible than vocal campaigns (Farooq and Wicaksono, 2021). Consumers can distinguish between honest green and greenwashed products more effectively when they are prompted to evaluate greenwashing cues explicitly, which underlines the importance of claim design and interpretation (Fella and Bausa, 2024). At the strategic level, sustainability contributes to competitiveness when it is embedded in products, stakeholder value creation, and innovation outcomes rather than treated as a symbolic add-on (Kahupi et al., 2021; Hermundsdottir and Aspelund, 2022). In food markets especially, labels and origin information act as credence cues that guide consumer choice when direct

verification of production attributes is difficult (Thøgersen, Pedersen and Aschemann-Witzel, 2019). Taken together, this literature suggests that customer-facing sustainability knowledge is most likely to generate competitive value when firms align three levers: claim integrity, choice guidance, and relational sensemaking.

3 Methodology

The paper adopts a qualitative multi-case design because the objective is to explain an emerging phenomenon and develop analytically generalisable mechanisms rather than test a mature theory (Yin, 2018). Theoretical sampling guided case selection so that each case could illuminate similarities and contrasts relevant to the research question, which is consistent with theory-building logic in case research (Eisenhardt, 1989; Eisenhardt and Graebner, 2007). The two industries were selected for three reasons. First, both food and fashion are highly visible to consumers and strongly exposed to sustainability-sensitive choice. Second, both industries combine digital touchpoints with human mediation. Third, they differ in the immediacy with which customers can verify sustainability-related outcomes, allowing stronger theoretical contrast between cases. The four focal cases were chosen as bounded examples of customer-facing sustainability knowledge pathways. McDonald's and H&M were selected as cases in which human-mediated communication was especially visible through influencer or campaign storytelling. Carrefour, in partnership with Too Good To Go, was selected as a case in which digital intermediation translated sustainability into a purchase architecture. SHEIN was selected as a contrasting case in which AI-mediated recommendation could amplify sustainability cues at scale while leaving verification contested. The unit of analysis is the customer-facing sustainability knowledge pathway through which a firm's sustainability-related information reaches customers and influences perceived competitiveness. This unit of analysis shifts attention from the firm in general to the specific combination of actors, artefacts, interfaces, and interpretive processes that shape customer judgement. The study combined primary and secondary data. Primary data consisted of anonymous semi-structured interviews of customers from the selected case studies. Secondary data included company reports, webpages, campaign materials, sustainability documents and platform descriptions. Triangulation across these sources increased the robustness of within-case interpretations (Yin, 2018). The analysis followed the Gioia methodology to move inductively from empirical material to a structured explanatory framework (Gioia, Corley and Hamilton, 2013). First-order concepts were generated from observed claims, interfaces, cues, explanations, and customer-facing interactions. These concepts were then grouped into second-order themes that captured how sustainability meaning was verified, translated, or socially interpreted. Finally, three aggregate dimensions emerged: claim integrity, choice guidance, and relational sensemaking. To answer the reviewers' concern regarding transparency, alignment across the three levers was assessed through a simple comparative rule. A case was treated as aligned on a lever when the customer journey displayed a consistent mechanism for that lever across the dominant touchpoints. Cases were then classified as two-lever, one-lever, or absent alignment configurations according to the number of levers clearly supported by the observed pathway. Case quality was strengthened through protocol-driven case documentation, within-case narratives, cross-case comparison, and explicit reflection on how interpretations were derived from the empirical material, in line with recommendations for rigorous case research (Goffin et al., 2019).

4 Findings

4.1 Claim integrity

Claim integrity captures whether sustainability claims appear auditable, standardised, and bounded by intelligible explanation. This dimension emerged when the coding process distinguished between cues that merely attracted attention and cues that allowed the customer to verify what the claim meant. In the Carrefour case, claim integrity was comparatively strong because the sustainability proposition was embedded in a concrete commercial mechanism. The discounted surprise basket made the anti-food-waste claim tangible at the moment of purchase. The customer did not need to infer a distant environmental virtue from a broad narrative; the value proposition was directly connected to the rescue of unsold food and to an observable price discount. Digital mediation therefore translated the sustainability claim into a bounded and testable commercial offer. In McDonald's and H&M, claim integrity was more fragile. Human-mediated storytelling increased salience and trust by attaching sustainability meaning to familiar brand narratives and socially credible messengers. Yet the same strength also introduced ambiguity. The sustainability message depended heavily on how influencers or campaign materials framed the offering. In these cases, human mediation made the claim understandable, but it did not always make its limits equally visible. As a result, customers could recognise the intended

sustainable positioning without receiving the same degree of auditable substantiation found in the Carrefour pathway. SHEIN represented the weakest configuration on this dimension. Here, sustainability cues appeared to benefit from algorithmic amplification and repeated exposure, but not from a robust verification architecture. AI-mediated visibility could raise the prominence of eco-related labels during browsing, yet the browsing journey offered little support for customers who wished to establish what made the highlighted item more sustainable in operational terms. In this configuration, AI acted mainly as an amplifier of eco-signals rather than as a verifier of claim content. Across cases, claim integrity was highest when AI standardised and bounded the claim while human channels clarified trade-offs and limits. Where human mediation operated without auditable cues, or AI visibility operated without verification, sustainability meaning became more contestable and more exposed to scepticism.

4.2 Choice guidance

Choice guidance refers to the extent to which sustainability information is translated into decision-ready comparisons at the point of selection. This dimension emerged from the distinction between cases that made sustainability actionable in the buying moment and cases that treated sustainability mainly as a background narrative. Carrefour again provided the clearest evidence of strong choice guidance. The digital interface and offer structure linked sustainability with price, availability, and immediate purchase action. The customer did not simply learn that waste reduction was desirable; the pathway guided the customer toward a specific decision in which sustainability and value converged. This is an example of hybrid-knowledge management at its most practical: information was not only communicated but converted into a simple, comparable, and economically legible choice. SHEIN also displayed a powerful form of choice guidance, but of a different kind. Recommendation routines, visibility ranking, and repeated eco-cues could channel attention toward items carrying sustainability labels. However, this guidance was stronger on navigation than on informed comparison. The customer was guided to click, view, and possibly purchase, but the rationale remained thinner than in the Carrefour case. The result was behavioural direction without equivalent interpretive depth. In McDonald's and H&M, choice guidance was weaker because the sustainability message was not always anchored in the exact moment of selection. Human endorsement could stimulate trial, brand warmth, or perceived relevance before the choice, but unless menus, labels, or filters reinforced the message at the final decision point, the customer still had to bridge the gap between general sustainability meaning and concrete selection criteria. The cross-case pattern therefore suggests that sustainable competitiveness benefits when AI converts sustainability claims into decision-ready comparisons and when human mediation ensures that the comparison remains intelligible rather than merely persuasive. Where this translation is absent, sustainability remains a favourable association but not a strong basis for customer choice.

4.3 Relational sensemaking

Relational sensemaking captures the extent to which ambiguity can be resolved through trusted interaction and sustained after the initial purchase decision. This dimension emerged as distinct from the other two because some cases were effective in explanation and reassurance even when their point-of-choice design remained weak. McDonald's and H&M showed the clearest signs of relational sensemaking. Human-mediated communication made sustainability easier to connect with everyday meanings, lifestyle aspirations, and socially shared language. In these cases, people rather than systems played the central role in reducing ambiguity. Human actors could translate abstract sustainability cues into narratives that customers recognised as relevant to eating habits, fashion identity, or responsible consumption. This did not fully solve the problem of verification, but it reduced interpretive distance and strengthened authenticity. Carrefour showed a more partial version of relational sensemaking. The digital pathway offered continuity after purchase by allowing the customer to connect the bargain with a concrete anti-waste outcome, yet the pathway was less relational in the interpersonal sense. The customer received continuity and coherence, but not the same capacity for trust repair, reassurance, or conversational explanation that human channels can provide. SHEIN again represented the weakest configuration. When ambiguity arises in a highly algorithmic environment, the customer has fewer relational resources with which to interpret or challenge sustainability cues. Without trusted human mediation, the pathway depends disproportionately on interface signals. This can sustain exposure, but it does not readily repair doubt once doubt is activated. Considering the three levers together, none of the cases displayed complete alignment. Carrefour approximated a two-lever configuration, with strong claim integrity and choice guidance but only partial relational sensemaking. McDonald's and H&M approximated one-to-two-lever configurations centred on human interpretation rather than auditable verification. SHEIN represented a one-lever configuration in which guidance through visibility was not matched by integrity or relational support. These

patterns explain why sustainable competitiveness was strengthened only partially across the observed cases: when one or more levers were missing, sustainability value remained visible but unstable.

5 Conclusions

The study set out to explain how sustainable competitiveness can be strengthened through hybrid-knowledge mechanisms involving both humans and AI agents. The evidence indicates that the key issue is not the simple presence of AI or human mediation, but the alignment of three customer-facing levers: claim integrity, choice guidance, and relational sensemaking. Hybrid-knowledge management becomes strategically valuable when firms govern these levers together across the customer journey. The paper makes three theoretical contributions. First, it extends knowledge-based and knowledge-creation perspectives by shifting attention from internal knowledge stocks to customer-facing sustainability knowledge flows as a source of competitive value (Grant, 1996; Nonaka and Toyama, 2003). Second, it refines the literature on human–AI partnership in knowledge management by showing that hybrid-knowledge management is not only a technological complementarity but also a governance capability that coordinates verification, translation, and interpretation across touchpoints (Lu, Nerode and Subrahmanian, 1996; Jarrahi et al., 2023). Third, it bridges CSR awareness and communication research with work on algorithmic transparency and trust by explaining why sustainability communication creates competitive value only when awareness becomes credible, decision-relevant, and interpretable to customers (Servaes and Tamayo, 2013; Wanner et al., 2022). Managerially, the findings imply that firms should not rely on one mediation path alone. AI systems should be designed to surface auditable cues, standardise labels, and make sustainability comparisons explainable at the moment of choice. Human channels should then be equipped to translate these cues, explain trade-offs, and repair ambiguity when customer doubt emerges. Different configurations require different managerial priorities. Where firms already possess strong digital guidance but weak trust, they need more human interpretation. Where they possess strong storytelling but weak substantiation, they need stronger verification infrastructures. For policymakers, the results support stricter comparability and traceability requirements across both digital and human channels. If recommendation systems elevate sustainability cues, there is a growing case for transparency duties regarding the basis of that elevation, especially when customers cannot otherwise distinguish verified information from persuasive amplification. The study has limitations. It relies on four cases and on publicly observable customer-facing pathways rather than on direct measurement of customer behaviour. In addition, no case displayed full alignment across all three levers, which means that the proposed ideal-type remains analytically inferred rather than fully observed. Future research could test the framework quantitatively, examine boundary conditions across industries and customer segments, and connect alignment configurations to outcomes such as willingness to pay, retention, or backlash risk. Longitudinal research could also explore how repeated exposure to hybrid mediation changes trust over time.

6 Ethics Declaration

The study did not raise any specific ethical concerns, and formal ethical approval was not required.

7 AI Declaration

The authors declare that no AI tools were used in the preparation, analysis, interpretation, or writing of this manuscript. The authors take full responsibility for the content of the paper.

References

Berger, B., Adam, M., Rühr, A. and Benlian, A. (2021) “Watch me improve—algorithm aversion and demonstrating the ability to learn”, Business & Information Systems Engineering, Vol. 63, No. 1, pp. 55-68.

Cambra-Fierro, J., Melero-Polo, I. and Vázquez-Carrasco, R. (2014) “The role of frontline employees in customer engagement”, Revista Española de Investigación de Marketing ESIC, Vol. 18, No. 2, pp. 67-77.

De Veirman, M., Cauberghe, V. and Hudders, L. (2017) “Marketing through Instagram influencers: The impact of number of followers and product divergence on brand attitude”, International Journal of Advertising, Vol. 36, No. 5, pp. 798-828.

Delmas, M.A. and Burbano, V.C. (2011) “The drivers of greenwashing”, California Management Review, Vol. 54, No. 1, pp. 64-87.

Dietvorst, B.J., Simmons, J.P. and Massey, C. (2015) “Algorithm aversion: People erroneously avoid algorithms after seeing them err”, Journal of Experimental Psychology: General, Vol. 144, No. 1, pp. 114-126.

Du, S., Bhattacharya, C.B. and Sen, S. (2010) “Maximizing business returns to corporate social responsibility (CSR): The role of CSR communication”, International Journal of Management Reviews, Vol. 12, No. 1, pp. 8-19.

Eisenhardt, K.M. (1989) “Building theories from case study research”, Academy of Management Review, Vol. 14, No. 4, pp. 532-550.
Eisenhardt, K.M. and Graebner, M.E. (2007) “Theory building from cases: Opportunities and challenges”, Academy of Management Journal, Vol. 50, No. 1, pp. 25-32.
Farooq, Y. and Wicaksono, H. (2021) “Advancing on the analysis of causes and consequences of green skepticism”, Journal of Cleaner Production, Vol. 320, 128927.
Fella, S. and Bausa, E. (2024) “Green or greenwashed? Examining consumers’ ability to identify greenwashing”, Journal of Environmental Psychology, Vol. 95, 102281.
Gioia, D.A., Corley, K.G. and Hamilton, A.L. (2013) “Seeking qualitative rigor in inductive research: Notes on the Gioia methodology”, Organizational Research Methods, Vol. 16, No. 1, pp. 15-31.
Goffin, K., Åhlström, P., Bianchi, M. and Richtnér, A. (2019) “State-of-the-art: The quality of case study research in innovation management”, Journal of Product Innovation Management, Vol. 36, No. 5, pp. 586-615.
Grant, R.M. (1996) “Toward a knowledge-based theory of the firm”, Strategic Management Journal, Vol. 17, No. S2, pp. 109-122.
Hermundsdottir, F. and Aspelund, A. (2022) “Competitive sustainable manufacturing: Sustainability strategies, environmental and social innovations, and their effects on firm performance”, Journal of Cleaner Production, Vol. 370, 133474.
Jarrahi, M.H., Askay, D., Eshraghi, A. and Smith, P. (2023) “Artificial intelligence and knowledge management: A partnership between human and AI”, Business Horizons, Vol. 66, No. 1, pp. 87-99.
Kahupi, I., Hull, C.E., Okorie, O. and Millette, S. (2021) “Building competitive advantage with sustainable products: A case study perspective of stakeholders”, Journal of Cleaner Production, Vol. 289, 125699.
Logg, J.M., Minson, J.A. and Moore, D.A. (2019) “Algorithm appreciation: People prefer algorithmic to human judgment”, Organizational Behavior and Human Decision Processes, Vol. 151, pp. 90-103.
Lou, C. and Yuan, S. (2019) “Influencer marketing: How message value and credibility affect consumer trust of branded content on social media”, Journal of Interactive Advertising, Vol. 19, No. 1, pp. 58-73.
Lu, J.J., Nerode, A. and Subrahmanian, V.S. (1996) “Hybrid knowledge bases”, IEEE Transactions on Knowledge and Data Engineering, Vol. 8, No. 5, pp. 773-785.
Nonaka, I. and Toyama, R. (2003) “The knowledge-creating theory revisited: Knowledge creation as a synthesizing process”, Knowledge Management Research & Practice, Vol. 1, No. 1, pp. 2-10.
Poszler, F. and Lange, B. (2024) “The impact of intelligent decision-support systems on humans’ ethical decision-making: A systematic literature review and an integrated framework”, Technological Forecasting and Social Change, Vol. 204, 123403.
Servaes, H. and Tamayo, A. (2013) “The impact of corporate social responsibility on firm value: The role of customer awareness”, Management Science, Vol. 59, No. 5, pp. 1045-1061.
Thøgersen, J., Pedersen, S. and Aschemann-Witzel, J. (2019) “The impact of organic certification and country of origin on consumer food choice in developed and emerging economies”, Food Quality and Preference, Vol. 72, pp. 10-30.
Wanner, J., Herm, L.-V., Heinrich, K. and Janiesch, C. (2022) “The effect of transparency and trust on intelligent system acceptance: Evidence from a user-based study”, Electronic Markets, Vol. 32, No. 4, pp. 2079-2102.
Yin, R.K. (2018) Case study research and applications, Sage, Thousand Oaks, CA.

The Key Factors in RTO Spin-off Development: Reconfiguring Organizational Knowledge for Spin-off Ventures

Guan-Ting Chen, Sirirat Sae Lim
Institute of Management of Technology, National Yang Ming Chiao Tung University, Hsinchu, Taiwan

Abstract

Research and Technology Organizations (RTOs) are under growing pressure to convert publicly funded knowledge assets into market-facing ventures, yet the mechanisms through which RTOs reconfigure knowledge across the parent-spin-off boundary remain insufficiently theorized. Drawing on the Knowledge-Based View, knowledge-creation theory, and intellectual-capital theory, this paper asks how RTOs combine intellectual property (IP), tangible assets, tacit know-how, human capital, and relational capital during spin-off development, and how Resource Coordination Capability enables such reconfiguration. The study uses a qualitative, theory-elaborating design based on semi-structured interviews with twenty-two respondents, including thirteen investors and nine policymakers with direct exposure to RTO spin-off cases. This external stakeholder sample was selected because investors and policymakers observe multiple reconfiguration attempts and evaluate whether ventures are investment-ready or policy-ready. A six-phase thematic analysis generated two linked theoretical artifacts. First, the Knowledge Asset Reconfiguration (KAR) Model shows that licensing, assignment, and joint development are normally combined across different subsets of the knowledge-asset portfolio rather than chosen as mutually exclusive alternatives. Second, the Progressive Separation Model explains spin-off emergence as a three-phase process of incubation, acceleration, and independent operation, with different capital forms moving at different speeds. Findings indicate that structural capital can be formalized relatively quickly, human capital migrates in clusters around formation, and relational capital remains the slowest and most consequential asset to rebuild. Across both models, Resource Coordination Capability acts as the meta-capability that allows RTOs to assemble, sequence, and reallocate heterogeneous resources across a changing boundary. The paper contributes to RTO spin-off research by replacing a single-transfer view with a boundary-reconfiguration view. This processual framing treats the venture not as an entity born at incorporation but as an evolving arrangement through which codified IP, people, routines, facilities, and relationships are progressively recombined. It is especially relevant for publicly funded technologies, where market uncertainty, policy accountability, and knowledge-leakage concerns must be balanced. The paper offers managers, investors, and policymakers practical diagnostics for judging reconfiguration readiness.

Keywords: Research and Technology Organization; spin-off; Knowledge-Based View; knowledge asset reconfiguration; Resource Coordination Capability.

Paper type: Academic Research Paper

1 Introduction

Public research organizations across Europe, Asia, and North America are increasingly expected to do more than produce publications and patents. Research and Technology Organizations (RTOs), including institutions such as Fraunhofer, ITRI, CSIRO, and SINTEF, are asked to translate publicly funded research into commercially viable spin-off ventures that create employment, tax revenue, industrial upgrading, and sovereign technology capability. The policy objective is clear. The organizational mechanics are less clear.

Much of the technology-transfer literature treats spin-off formation as a discrete decision: license the technology, assign the intellectual property, or create a joint venture. Each modality is often compared on control, revenue capture, and transaction cost. Practitioners describe a more complicated process. A patent may be assigned at formation, while process know-how remains inside the RTO, a laboratory facility is shared, and future improvements are governed by joint-development clauses. Recent work on academic spin-offs has therefore begun to shift attention from one-off transfer events to organization-embedded emergence (Opizzi, Loi, and Lamine, 2025; Prokop, 2023).

The stakes of this distinction are practical as well as theoretical. Public sponsors want evidence that publicly funded knowledge is being mobilized rather than privatized too cheaply. Investors want evidence that a venture controls enough knowledge to execute, but also enough access to the parent organization to learn. RTO managers must protect the scientific platform while giving the venture room to build a distinct market identity. These goals cannot be reconciled through a single legal transaction.

This empirical pattern creates a theoretical gap. The Knowledge-Based View (KBV) treats knowledge as the strategically central resource of the firm and emphasizes the coordination mechanisms required to integrate specialized knowledge (Grant, 1996; Kogut and Zander, 1992). However, the RTO-spin-off interface is not a stable

firm. It is a boundary in motion, where some knowledge must move, some must remain accessible, and some must be jointly created after separation. Complementary perspectives on tacit-explicit knowledge conversion (Nonaka and Takeuchi, 1995) and intellectual capital (Bontis, 1998) help specify what moves, but they have rarely been integrated into a process account of RTO spin-off development.

The study asks two research questions. RQ1: Through what configurations of IP modalities and tangible or intangible assets do RTOs reconfigure organizational knowledge across the parent-spin-off boundary? RQ2: How does Resource Coordination Capability shape whether these configurations produce investable ventures? The paper contributes by developing the Knowledge Asset Reconfiguration (KAR) Model, by specifying a Progressive Separation Model of incubation, acceleration, and independent spin-off, and by using investor and policymaker perspectives to reveal diagnostic criteria that are often hidden in internal RTO accounts.

2 Theoretical Background

2.1 Knowledge-Based View and Boundary Reconfiguration

The KBV refines the resource-based view by treating knowledge as the firm's most important productive input. Grant (1996) argues that knowledge is largely held by individuals, that organizations exist to integrate specialized knowledge, and that integration depends on mechanisms such as rules, routines, sequencing, and group problem solving. Later reviews confirm that KBV remains influential in international business and capability research, particularly where knowledge must be recombined under uncertainty (Stoian, Tardios, and Samdanis, 2024; Tamirat and Amentie, 2023).

For RTO spin-offs, these premises have three implications. First, transferring a patent does not transfer all knowledge required to use the patent; tacit know-how remains embedded in scientists, engineers, laboratory routines, and accumulated troubleshooting experience. Second, integration mechanisms that worked inside the RTO may fail once founders, legal rights, equipment, and customer relationships are split across organizations. Third, a spin-off is not only a recipient of knowledge but also a new site of knowledge integration. Spin-off development is therefore a boundary-reconfiguration problem rather than a simple boundary-crossing event.

This boundary-reconfiguration view also clarifies why spin-off timing is difficult. If separation occurs too late, the venture may remain dependent on the parent and lose entrepreneurial momentum. If separation occurs too early, the venture may acquire formal rights without the routines required to use them. The challenge is not merely to transfer knowledge but to design a sequence in which knowledge can continue to be integrated while organizational autonomy increases.

2.2 Knowledge Creation and Intellectual Capital

Nonaka and Takeuchi's (1995) SECI model explains knowledge creation through socialization, externalization, combination, and internalization. In the RTO context, externalization is visible when tacit laboratory knowledge is codified into patents, specifications, and manuals. Combination occurs when the spin-off recombines transferred IP with external technologies. Socialization is less visible but critical, because founders and technical staff carry process knowledge through shared work. Internalization begins when the venture converts licensed or assigned knowledge into its own operational routines.

Bontis's (1998) taxonomy of intellectual capital further decomposes the asset base into human, structural, and relational capital. Human capital includes individual expertise and experience; structural capital includes patents, data, routines, software, and organizational procedures; relational capital includes customer, supplier, regulator, and partner relationships. Cosa, de Matos Pedro, and Urban (2024) show that this tripartite logic remains central to contemporary intellectual-capital measurement. In spin-off formation the three capital forms rarely move together. Structural capital may be codified and transferred quickly, human capital moves when people change roles, and relational capital must be rebuilt through repeated external interactions.

Combining SECI with intellectual-capital theory allows the analysis to avoid an IP-only account. A patent is structural capital and an outcome of externalization, but the ability to enact the patent may depend on human capital acquired through socialization. A first customer introduction is relational capital, but the venture must internalize the relationship by becoming credible in its own right. The same technology therefore contains different knowledge assets that move through different mechanisms.

2.3 IP Modalities and Resource Coordination Capability

RTOs commonly use three IP modalities. Licensing preserves parent ownership while granting the spin-off defined rights of use. Assignment transfers ownership, often in return for equity, payment, or a strategic public-

interest arrangement. Joint development defines how future inventions will be owned or used when the parent and spin-off continue to collaborate. Prior research often treats these modalities as substitutes (Clarysse et al., 2005; Grimaldi et al., 2011; Rothaermel, Agung, and Jiang, 2007). This paper instead treats them as instruments that can be combined across different assets and phases.

Resource Coordination Capability refers to the RTO's capacity to assemble, sequence, and reallocate human, financial, infrastructural, legal, and relational resources across organizational boundaries over time. It is not simply resource ownership or technical excellence. It is the capability to decide which asset moves under which modality, when people should migrate, how shared infrastructure should be governed, and how external relationships should be transferred or substituted. This construct connects KBV's integration logic with the operational reality of spin-off formation.

The capability is visible in mundane organizational practices. High-capability RTOs maintain clear decision rights between technology-transfer offices, legal units, laboratory directors, and executive sponsors. They can distinguish assets that should be transferred from assets that should remain accessible through collaboration. Low-capability RTOs tend to convert uncertainty into delay, or to force diverse assets into one contract because the organization lacks a process for managing a hybrid arrangement.

3 Conceptual Framework and Proposed Models

3.1 Knowledge Asset Reconfiguration Model

The KAR Model treats RTO spin-off development as a portfolio problem. The relevant question is not whether the RTO should license, assign, or jointly develop IP. The question is which subsets of the knowledge-asset portfolio should be governed by each modality at a specific moment. A core platform patent may be assigned, background technology may remain licensed, and next-generation improvements may be placed under joint development. Tangible assets, such as pilot equipment, can follow a different route from intangible assets, such as process know-how or regulatory credibility.

The model therefore shifts attention from choosing the best modality to maintaining fit between asset characteristics and governance arrangements. Assets that are mature, codified, and separable may be assigned. Assets that are valuable but still connected to the parent platform may be licensed. Assets that will evolve through continued collaboration require joint-development governance. A venture may need all three at once if it is to satisfy investor demands for control while preserving access to parent knowledge.

Three propositions follow. P1: RTO spin-offs using hybrid IP configurations will be perceived by investors as more reconfiguration-ready than those using a single modality. P2: Asset portfolios rich in tacit know-how benefit more from joint-development anchors, because these preserve socialization channels. P3: Resource Coordination Capability is a necessary condition for executing hybrid configurations without creating legal, operational, or relational fragmentation.

Table 1: KAR Model: IP Modalities and Reconfiguration Challenges

IP modality	Knowledge-asset role	Capital and SECI emphasis	Coordination challenge
Licensing	Preserves parent ownership while granting defined use rights.	Structural capital; combination of explicit knowledge.	Scope control, royalty monitoring, and continued access to background know-how.
Assignment	Transfers ownership of selected platform or application-specific IP.	Structural capital with human-capital support; externalization and internalization.	Avoiding tacit-knowledge loss and preserving founder-team continuity.
Joint development	Governs future inventions and ongoing parent-spin-off collaboration.	Structural and relational capital; socialization and externalization.	Allocating future IP, aligning incentives, and sustaining knowledge channels.

3.2 Progressive Separation Model

The Progressive Separation Model explains the temporal side of reconfiguration. In incubation, the prospective venture remains embedded in the RTO and learns through shared routines. In acceleration, selective decoupling begins: patents are formalized, some assets are assigned, human capital starts to migrate, and joint-development rules are negotiated. In the spin-off phase, the venture becomes legally and operationally independent but remains relationally coupled to the parent while it builds its own market standing.

Separation is progressive because legal independence, operational independence, and relational independence do not arrive at the same time. Incorporation may create a legal boundary before routines have stabilized. Assignment may transfer ownership before the venture has full operating confidence. Customer introductions may continue long after the venture has its own employees and accounts.

The model generates three further propositions. P4: Ventures that compress incubation are more likely to fail because they lack absorptive capacity. P5: Acceleration is the highest-risk phase because IP, governance, and human-capital reconfiguration occur concurrently. P6: In the independent spin-off phase, relational-capital substitution becomes the binding constraint on performance.

Table 2: Progressive Separation Model

Phase	Knowledge position	Dominant transition	Main coordination task
Incubation	Venture remains embedded in the RTO.	Tacit learning and absorptive-capacity building.	Shield the venture from premature market exposure.
Acceleration	Selective decoupling becomes visible.	IP formalization, partial assignment, founder migration, and joint-development governance.	Time legal, human, and governance changes without destabilizing either organization.
Spin-off	Venture is independent but relationally coupled.	Relational-capital substitution and autonomous routine building.	Replace parent-derived credibility with independent market standing.

4 Methodology

4.1 Research Design and Sample

The study adopts a qualitative, theory-elaborating design. Existing theory supplied the sensitizing concepts: KBV for knowledge integration, SECI for tacit-explicit conversion, intellectual capital for asset categories, and Resource Coordination Capability for cross-boundary orchestration. Empirical analysis then elaborated these concepts into the KAR and Progressive Separation Models.

The sample consisted of twenty-two respondents: thirteen investors and nine policymakers. Investors included venture capital partners, corporate venture investors, and government-affiliated investment-fund managers. Policymakers included officials from technology-policy ministries, agency leaders responsible for RTO oversight, and public program directors. All respondents had direct decision-making exposure to at least three RTO spin-off cases within the preceding five years. The sample intentionally privileges an external vantage point because investors and policymakers compare multiple reconfiguration attempts and must judge whether a spin-off is more than a technology demonstration.

This sampling strategy also reduces reliance on self-reported organizational success. RTO managers and founders may emphasize strategic intent, scientific promise, or entrepreneurial commitment. External stakeholders instead ask whether asset control, team migration, governance, and market relationships are credible enough to justify capital or policy support.

4.2 Data Collection and Analysis

Semi-structured interviews of sixty to ninety minutes were conducted between September 2025 and February 2026 in Mandarin Chinese or English. The protocol moved from general framing of RTO spin-offs to case-specific discussion of IP modality, tangible and intangible asset movement, phase-specific risks, and diagnostic cues for reconfiguration readiness. Interviews were recorded with consent, transcribed, anonymized, and analyzed using a six-phase thematic-analysis procedure (Braun and Clarke, 2006, 2019).

Coding combined deductive and inductive steps. Initial coding in NVivo 14 produced 147 first-order codes. Axial coding consolidated these into 24 second-order themes under six aggregate dimensions: IP modality configuration, tangible-asset reconfiguration, intangible-asset reconfiguration, phase-specific coordination challenges, Resource Coordination Capability indicators, and investor or policymaker diagnostic cues. A 25 percent sub-sample was double-coded by an independent coder, producing Cohen's kappa of 0.81. Trustworthiness was strengthened through member checking with six respondents, triangulation with public case documentation, and an audit trail of coding decisions.

The analysis treated respondent quotations as evidence of recurring evaluation logics rather than as isolated opinions. Themes were retained only when they appeared across respondent types or when a minority pattern

explained a theoretically important exception, such as policy regimes that restricted IP assignment but still produced spin-offs through joint-development vehicles.

5 Findings

5.1 Hybrid IP Configurations

The most consistent empirical pattern was the combined use of licensing, assignment, and joint development. Respondents rarely described successful spin-offs as clean transfers. Instead, they expected a structured bundle in which different assets moved under different instruments. One investor summarized the diagnostic logic: "A mature RTO gives you a portfolio: a core assignment, a platform license, and a joint-development clause for the next generation. That tells me somebody thought about it."

Investors treated single-modality deals as warning signs because they often concealed unaddressed tacit knowledge, data, or future improvement rights. Policymakers made a similar point in regulatory language. When rules restricted outright assignment of publicly funded IP, spin-off activity did not necessarily decline; it shifted toward licensing and joint-development vehicles. This supports the KAR Model's claim that policy constraints change the modality mix rather than eliminating reconfiguration, provided coordination capability is strong enough.

Two sub-patterns appeared repeatedly. First, respondents distinguished platform IP from application-specific IP. Platform IP was often licensed to preserve public or parent optionality, while application-specific improvements were more likely to be assigned to the venture. Second, joint-development clauses functioned as anchors for future knowledge creation. Even where pre-existing IP was assigned, respondents expected a mechanism for handling improvements, data, and derivative inventions generated after launch.

5.2 Tangible and Intangible Asset Reconfiguration

Respondents distinguished sharply between tangible assets and intangible assets. Tangible assets, such as equipment, prototype inventories, and pilot infrastructure, were described as administratively heavy but conceptually straightforward. Intangible assets, especially process know-how and relational capital, were described as the real bottleneck. As one investor put it, "The equipment transfer is a purchase order. The know-how transfer is a human-resources problem. The relational-capital transfer is a three-year project."

The three intellectual-capital forms moved at different speeds. Structural capital reconfigured fastest once legal instruments were ready. Human capital moved in clusters as founders and engineers shifted time allocation or employment. Relational capital moved slowest because customers, regulators, suppliers, and strategic partners had to learn to trust the spin-off without relying entirely on the RTO's reputation. Respondents repeatedly identified this relational substitution as the determinant of survival in the two-to-five-year period after formation.

This finding explains why formal asset lists often overstate readiness. A venture may show assigned patents, transferred equipment, and founder employment agreements while still lacking the routines and relationships that make the technology usable. Investors therefore examined evidence of operational learning: whether engineers had run the process outside the parent laboratory, whether customers had interacted directly with the venture, and whether regulatory bodies recognized the venture as a credible counterparty.

5.3 Phased Dynamics of Progressive Separation

Interview accounts supported the three-phase separation model. Incubation was often invisible to external stakeholders but crucial to absorptive capacity. Respondents argued that ventures pushed out too early may possess formal IP but lack the tacit routines needed to operate it. Acceleration was the phase where tensions became visible, because legal transfer, founder migration, infrastructure access, and governance of future IP had to be resolved at the same time.

Acceleration was also the phase in which misalignment became expensive. If the legal team optimized assignment terms without understanding technical dependency, the venture could lose access to necessary know-how. If the laboratory delayed founder release, the venture could miss investor milestones. If policymakers pushed for rapid commercialization without supporting governance of future IP, the spin-off could inherit unresolved conflicts.

In the spin-off phase, formal independence did not mean complete separation. Ventures continued to depend on parent-derived credibility, introductions, test facilities, and regulatory history. One investor noted that a venture may call itself independent while its customer list still reveals dependence on parent

introductions. Technical and financial indicators could appear healthy twelve to eighteen months after formation, while relational indicators lagged by several more years.

5.4 Resource Coordination Capability

Across both models, Resource Coordination Capability was the strongest discriminator between successful and unsuccessful reconfiguration. Respondents described it through observable cues: clear decision rights among the technology-transfer office, legal counsel, scientific leadership, and executive sponsors; routinized processes for resolving trade-offs; and a documented history of prior spin-offs or licensing arrangements. High-capability RTOs were able to adapt the mix of licensing, assignment, and joint development as new information emerged.

Low-capability RTOs tended to simplify prematurely. They used a single legal instrument because it was administratively convenient, delayed decisions about founder migration, or relied too heavily on the parent organization's reputation after launch. The resulting ventures could look complete on paper while remaining fragile in practice. Resource Coordination Capability therefore operates as a meta-capability: it does not replace scientific excellence, IP quality, or market demand, but it governs whether those elements can be assembled into an investable venture.

Respondents emphasized that the capability can be assessed early. Meetings in which legal, technical, and commercial actors could discuss trade-offs without escalation were interpreted as positive signals. Meetings in which each actor deferred to another unit were interpreted as evidence that the RTO had not developed routines for cross-boundary orchestration.

6 Discussion

The findings extend KBV by specifying a class of knowledge-integration problems in which the boundary of integration is itself under construction. Grant's (1996) mechanisms of rules, routines, sequencing, and group problem solving remain central, but they must operate across an emerging parent-spin-off boundary. Licensing preserves a coordination channel without moving ownership. Assignment moves codified structural capital but not the full tacit base. Joint development keeps a channel open for knowledge that will be created after separation.

The findings also refine SECI theory. Single assignment executes externalization and combination, but it does not ensure socialization. Joint-development arrangements can be interpreted as institutionalized socialization channels after formal separation. Incubation is socialization-heavy, acceleration is externalization- and combination-heavy, and the independent spin-off phase begins a new cycle of internalization as the venture converts transferred knowledge into its own routines.

Finally, the findings extend intellectual-capital theory by showing that human, structural, and relational capital do not reconfigure at the same speed. Structural capital moves through documents and contracts; human capital moves through career decisions and role redesign; relational capital moves through repeated external validation. Resource Coordination Capability is thus a form of structural capital that governs the movement of other capital forms.

Together, the two models reposition spin-off formation as a designed transition rather than a transaction. The KAR Model explains the content of the transition, while the Progressive Separation Model explains its timing. Their integration helps explain why ventures with strong technology may still fail: the formal knowledge asset may have moved, but the sequence of socialization, externalization, combination, internalization, and relational substitution may be incomplete.

7 Implications, Limitations, and Conclusion

7.1 Implications

For RTO managers, the KAR Model provides a practical diagnostic question: which assets should be licensed, assigned, shared, or jointly developed, and does the organization have the coordination capability to manage that mix? For policymakers, the Progressive Separation Model suggests that instruments should differ by phase. Incubation policies should protect learning; acceleration policies should support co-ownership governance and founder transition; spin-off policies should help ventures replace parent-derived relational capital with independent market standing.

For investors, Resource Coordination Capability offers due-diligence signals that complement technology and team assessment. Investors can examine whether the RTO's legal, technical, and executive actors make decisions

coherently, whether prior reconfigurations are documented, and whether the venture has a credible plan for relational-capital substitution.

The implication is not that every spin-off requires a complex contract. Rather, complexity should match the knowledge asset. Simple assignments may be appropriate for mature, separable assets. Hybrid configurations are needed when the venture depends on continuing access to parent routines, facilities, or external relationships.

7.2 Limitations and Future Research

The study has three limitations. First, the investor and policymaker sample illuminates cross-boundary evaluation but under-represents RTO managers and founders. Future research should test the models through multi-informant designs. Second, the qualitative design supports theory elaboration rather than statistical generalization. Quantitative research could test the propositions using survey or archival data. Third, the empirical base is concentrated in Asian-Pacific and European RTO ecosystems; comparative work should examine whether the models travel to U.S. federal laboratories and emerging-economy research institutes.

Future research could also examine ecosystem-level effects. If Resource Coordination Capability is unevenly distributed across RTOs, national spin-off output may depend not only on science quality or funding volume but also on the institutional routines that allow knowledge assets to be recombined across boundaries.

7.3 Conclusion

RTO spin-off development is not a single transfer decision. It is a boundary-reconfiguration process in which IP, equipment, tacit know-how, human capital, and relational capital move at different speeds under different instruments. The KAR Model explains the combinatorial content of this process, while the Progressive Separation Model explains its temporal phasing. Resource Coordination Capability determines whether the reconfiguration can produce ventures that are investable, governable, and capable of building independent market standing.

8 Ethics Declaration

The study was based on expert interviews with informed consent. Respondents were anonymized, and no personally identifying information is reported. The authors should insert any institutional ethics approval reference required by their university before final submission.

9 AI Declaration

Generative AI assistance (Gemini/ChatGPT) was used to support language condensation, editorial rewriting, and Word-format preparation. The authors reviewed the resulting manuscript and remain responsible for its content, arguments, citations, and declarations.

References

Bontis, N. (1998) Intellectual capital: an exploratory study that develops measures and models, Management Decision, Vol. 36, No. 2, pp. 63-76. https://doi.org/10.1108/00251749810204142

Braun, V. and Clarke, V. (2006) Using thematic analysis in psychology, Qualitative Research in Psychology, Vol. 3, No. 2, pp. 77-101. https://doi.org/10.1191/1478088706qp063oa

Braun, V. and Clarke, V. (2019) Reflecting on reflexive thematic analysis, Qualitative Research in Sport, Exercise and Health, Vol. 11, No. 4, pp. 589-597. https://doi.org/10.1080/2159676X.2019.1628806

Clarysse, B., Wright, M., Lockett, A., Van de Velde, E. and Vohora, A. (2005) Spinning out new ventures: a typology of incubation strategies from European research institutions, Journal of Business Venturing, Vol. 20, No. 2, pp. 183-216. https://doi.org/10.1016/j.jbusvent.2003.12.004

Cosa, M., de Matos Pedro, E. and Urban, B. (2024) How to assess the intellectual capital of firms in uncertain times: a systematic literature review and a proposed model for practical adoption, Journal of Intellectual Capital, Vol. 25, No. 7, pp. 1-29. https://doi.org/10.1108/JIC-02-2024-0046

Grant, R.M. (1996) Toward a knowledge-based theory of the firm, Strategic Management Journal, Vol. 17, No. S2, pp. 109-122. https://doi.org/10.1002/smj.4250171110

Grimaldi, R., Kenney, M., Siegel, D.S. and Wright, M. (2011) 30 years after Bayh-Dole: reassessing academic entrepreneurship, Research Policy, Vol. 40, No. 8, pp. 1045-1057. https://doi.org/10.1016/j.respol.2011.04.005

Hamilton, C. and Philbin, S.P. (2020) Knowledge based view of university tech transfer: a systematic literature review and meta-analysis, Administrative Sciences, Vol. 10, No. 3, 62. https://doi.org/10.3390/admsci10030062

Kogut, B. and Zander, U. (1992) Knowledge of the firm, combinative capabilities, and the replication of technology, Organization Science, Vol. 3, No. 3, pp. 383-397. https://doi.org/10.1287/orsc.3.3.383

Nonaka, I. and Takeuchi, H. (1995) The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation, Oxford University Press, New York.

Opizzi, M., Loi, M. and Lamine, W. (2025) Academic spin-off emergence as an organization-embedded dynamic: synthesis and research questions from a systematic literature review, Technovation, Vol. 146, Article 103292. https://doi.org/10.1016/j.technovation.2025.103292

Prokop, D. (2023) The academic spinoff theory of the firm, International Journal of Entrepreneurship and Innovation, Vol. 24, No. 4, pp. 233-243. https://doi.org/10.1177/14657503211066013

Rothaermel, F.T., Agung, S.D. and Jiang, L. (2007) University entrepreneurship: a taxonomy of the literature, Industrial and Corporate Change, Vol. 16, No. 4, pp. 691-791. https://doi.org/10.1093/icc/dtm023

Stoian, M.-C., Tardios, J.A. and Samdanis, M. (2024) The knowledge-based view in international business: a systematic review of the literature and future research directions, International Business Review, Vol. 33, No. 2, 102239. https://doi.org/10.1016/j.ibusrev.2023.102239

Tamirat, S. and Amentie, C. (2023) Advances in knowledge-based dynamic capabilities: a systematic review of foundations and determinants in recent literature, Cogent Business & Management, Vol. 10, No. 3, 2257866. https://doi.org/10.1080/23311975.2023.2257866

The Impact of Networking Skills on Quiet Quitting and Knowledge Hiding: Mediating Role of Job Crafting and Moderating Effect of Mindful Leadership

Wan-Jung Chung, Guo-Jia Bo, Hao-Chieh Lin
National Sun Yat-sen University, Kaohsiung City, Taiwan, R.O.C.

Abstract

As businesses recover from the impact of the COVID-19 pandemic, the post-pandemic workplace has undergone substantial transformation. Currently, many studies consider quiet quitting to be indicators of low work attitude, decreased work engagement, and reduced performance and retention willingness. By enabling employees to rediscover the value and meaning of their work through job reinvention, they can achieve work goals more energetically and efficiently (Lu & Chen, 2020). Connelly et al. (2012) argue that employees' knowledge-hiding behavior stems primarily from their distrust of the organization or their own interpersonal relationships within it (Černe et al., 2014; Connelly et al., 2012). When employees quietly leave, the organization may lose not only enthusiasm and additional investment, but also a weakening of knowledge flow, knowledge contributions, and collaborative wisdom.

From a knowledge management perspective, preventing quiet quitting is not only about work engagement, but also about the continuity of knowledge flow, knowledge sharing, and the operation of intelligent organizations. Cultivating trust can indeed help employees overcome self-interest and share knowledge, especially at the group level (Weng & Li, 2015). Although research on quiet quitting is increasing, existing literature mainly focuses on the results; comparatively, there is insufficient attention paid to how to prevent quiet quitting. To fill this gap, this study uses Conservation of Resources theory (COR) as its theoretical basis to explore how network capabilities as personal resources and job crafting as a proactive resource management behaviour affect quiet quitting. Furthermore, the discussion explores how a quiet quitting can affect knowledge hiding. Furthermore, this study also discusses the impact of mindfulness leadership on the process of transforming personal resources into proactive action.

To fill this gap, it examines networking ability as a personal resource and job crafting as a proactive form of resource management and investigates how these factors influence quiet quitting. It further explores how quiet quitting affects knowledge hiding. In addition, mindful leadership is conceptualized as a contextual resource that may shape the process through which personal resources are translated into proactive action.

Accordingly, this study investigates how networking ability influences quiet quitting through job crafting and how quiet quitting, in turn, affects knowledge hiding. Through this design, quiet quitting is positioned not merely as a problem of low engagement, but as a knowledge management risk with important organizational implications.

Keywords: Networking Ability, Job Crafting, Quiet Quitting, Mindful Leadership, Conservation of Resources (COR) Theory

Paper type: Academic Research Paper

1 Introduction

The post-pandemic workplace has undergone substantial transformation. Flexible work arrangements, accelerated digitalization, and the widespread use of AI systems have changed not only how work is performed, but also how employees relate to their organizations, roles, and colleagues. From a knowledge management perspective, preventing quiet quitting is not only about work engagement, but also about the continuity of knowledge flow, knowledge sharing, and the operation of intelligent organizations. Cultivating trust can indeed help employees overcome self-interest and share knowledge, especially at the group level (Weng & Li, 2015). Although research on quiet quitting is increasing, existing literature mainly focuses on the results; comparatively, there is insufficient attention paid to how to prevent quiet quitting. To fill this gap, this study uses Conservation of Resources theory (COR) as its theoretical basis to explore how network abilities as personal resources and job crafting as a proactive resource management behaviour affect quiet quitting. Furthermore, the discussion explores how a quiet quitting can affect knowledge hiding. Furthermore, this study also discusses the impact of mindfulness leadership on the process of transforming personal resources into proactive action.

To fill this gap, it examines networking ability as a personal resource and job crafting as a proactive form of resource management and investigates how these factors influence quiet quitting. It further explores how quiet quitting affects knowledge hiding. In addition, mindful leadership is conceptualized as a contextual resource that may shape the process through which personal resources are translated into proactive action.

Accordingly, this study investigates how networking ability influences quiet quitting through job crafting and how quiet quitting, in turn, affects knowledge hiding. Through this design, quiet quitting is positioned as a knowledge management risk with important organizational implications.

2 Literature Review and Hypothesis Development

2.1 Quiet quitting

That is often defined as "merely completing work tasks rather than investing in learning new skills or enhancing abilities.", The employee did not actually leave the company, but their work engagement and performance significantly decreased, and they were only fulfilling the basic responsibilities outlined in the job crafting（Scheyett, 2022）. Current research on quiet quitting largely focuses on work attitudes and performance consequences. Employees also realize that work is just work, work to earn a living, and once they leave their positions, the knowledge they have acquired becomes personal (Kao & Fu, 2023). However, employees' willingness and participation in knowledge sharing are still influenced by their intrinsic motivation (Meng-Chuan Tsai et al., 2018). Knowledge silencing initially exists only between two individuals but can erode trust mechanisms within an organization through continuous cycles, reducing organizational and individual creativity and performance, and ultimately potentially leading to the organization's overall collapse (Hung-Yu Chen & Hao-Chieh Lin., 2023).

Therefore, this study argues that quiet quitting may weaken the flow of tacit knowledge, impede collaborative problem-solving, and hinder the continuous updating of organizational knowledge, and should be understood as a knowledge management risk.

2.2 Networking ability and job crafting

According to the Conservation of Resources Theory (COR), individuals with abundant resources have a stronger motivation to "invest in resources," thereby initiating positive gain spirals to acquire more resources and mitigate potential losses (Hobfoll, 2011; Hobfoll et al., 2018). Within this framework, networking ability is considered a key personal trait resource, reflecting an individual's comprehensive effectiveness in building, maintaining, and utilizing interpersonal networks to obtain social support, critical information, and career influence. Employees with high networking ability excel at mobilizing external assistance, identifying environmental opportunities, and effectively transforming social capital into tangible professional knowledge.

Job crafting refers to the proactive structural adjustments employees make to task boundaries, interpersonal interactions, or cognitive frameworks to optimize personal-environment fit (P-E Fit) and the meaning of their work (Tims & Bakker, 2010; Wrzesniewski & Dutton, 2001). From a COR perspective, the information flow and interpersonal support brought by Networking Ability provide employees with the "psychological and social ammunition" needed for job crafting. When individuals have sufficient resources, they are more willing to invest resources (such as energy and social capital) to proactively rewrite their work scope, thereby achieving resource value-added. According to the Conservation of Resources Theory (COR), individuals with abundant resources have a stronger motivation to "invest in resources," thereby initiating positive gain spirals to acquire more resources and prevent potential losses (Hobfoll, 2011; Hobfoll et al., 2018). Based on this, this study proposes the following hypothesis:

H1: An individual's Networking Ability have a significant positive impact on their job crafting behaviour.

2.3 Job crafting and quiet quitting

When employees face persistently high work demands, role ambiguity, and resource scarcity, and their invested effort fails to yield commensurate resource returns, they face the threat of a "loss spiral" (Hobfoll, 2001; Hobfoll et al., 2018). To prevent further resource depletion, employees adopt defensive protective strategies, selectively reducing task engagement and emotional connection. In the contemporary context, quiet quitting can be seen as a psychological "withdrawal strategy" or resource-preserving behavior, aimed at maintaining remaining personal resources to avoid collapse (Klotz & Bolino, 2022).

In contrast, job crafting is a proactive resource-enhancing behaviour. Employees actively adjust their job design to improve their job fit by increasing structural or social resources, challenging requirements, or reducing hindering requirements (Tims et al., 2012; Bakker et al., 2012). From a COR perspective, job crafting can enhancing a sense of control and meaning in work, thereby suppressing the motivation for psychological withdrawal and negative engagement. Therefore, employees exhibiting higher levels of job crafting behaviour

have more abundant psychological resources and are less likely to need to protect themselves through quiet quitting. Based on this, this study proposes:

H2: There is a significant negative correlation between employees' level of job crafting and their tendency to leave quietly.

2.4 Quiet quitting and knowledge hiding

Knowledge hiding is defined as the intentional withholding or concealment of relevant information by an individual when others explicitly request it (Connelly et al., 2012). For organizations seeking innovation, fostering knowledge sharing among employees through institutional and psychological means is crucial for enhancing competitiveness (Jiang Dingyu & Lin Jiahua, 2019). According to Conservation of Resource Theory (COR), knowledge is considered a core personal resource with a competitive advantage, and sharing requires time, cognitive effort, and risk-taking. When employees perceive resource depletion or organizational reciprocity imbalances, they often employ hiding behavior as a defensive strategy to protect remaining resources and maintain career adaptability (Hobfoll et al., 2018; Xiao & Cooke, 2019).

When individuals perceive their resources as threatened or lost, they will mobilize existing resources to minimize the loss (Liu & Lin, 2022). Because the transformation and transmission of tacit knowledge highly depend on informal interactions and spontaneous input, this study argues that knowledge hiding is a psychological mechanism similar to quiet quitting.

In today's rapidly evolving knowledge economy, a large part of an organization's competitiveness comes from the knowledge possessed by its employees. For organizations that wish to innovate, how to promote knowledge sharing among employees through systems and mindset will be a key factor in enhancing their competitiveness (Jiang Dingyu & Lin Jiahua, 2019). When employees perceive resource depletion or organizational reciprocity imbalances, they often resort to covert behaviors as a defensive strategy to protect remaining resources and maintain personal career adaptability (Hobfoll et al., 2018; Xiao & Cooke, 2019).

If quiet quitting result in the loss of high-quality knowledge, it will directly weaken the organization's learning foundation and the digitization process of organizational wisdom (Cacciotti et al., 2020). Therefore, this study argues that quiet quitting are a key antecedent driving knowledge hiding and proposes the following hypothesis:

H3: Quiet quitting are positively correlated with knowledge hiding.

2.5 The mediating role of job crafting

According to the Conservation of Resources (COR) theory, owning resources is not the same as solving the problem. Individuals must increase the value of resources and resist the spiral of loss through "resource investment" (Hobfoll et al., 2018).

Employees with strong networking skills have greater "resource availability," which gives them a greater ability to reshape their work—for example, by obtaining feedback through interpersonal resources and by adjusting task boundaries through information advantages (Tims et al., 2012). In other words, network capabilities provide the "raw materials" for reshaping work, while work reshaping is the process of refining the raw materials into a "stress shield" (Halbesleben et al., 2014). This study argues that while networking provides employees with abundant resources such as information and social support, these external resources need to be actively internalized and transformed through job reengineering to be effective. Based on this, this study proposes:

H4: Job crafting mediates the relationship between networking ability and quiet quitting.

2.6 The moderating role of mindful leadership

Mindful leadership refers to the present-moment awareness, openness, and non-judgmental qualities that leaders demonstrate in their interactions with subordinates (Good et al., 2016). From the perspective of Conservation of Resource Theory (COR), mindful leadership not only represents external resources available to subordinates but also constitutes a stable "resource caravan passageway, " creating a psychologically safe and supportive environment (Hobfoll et al., 2018). If leaders lack mindfulness, even employees with networking resources may limit their initiative due to concerns about assessing risks or environmental pressures (Roche et al., 2014). Based on this, this study proposes:

H5: Mindful leadership has a positive moderating effect on the relationship between Networking Ability and Job crafting.

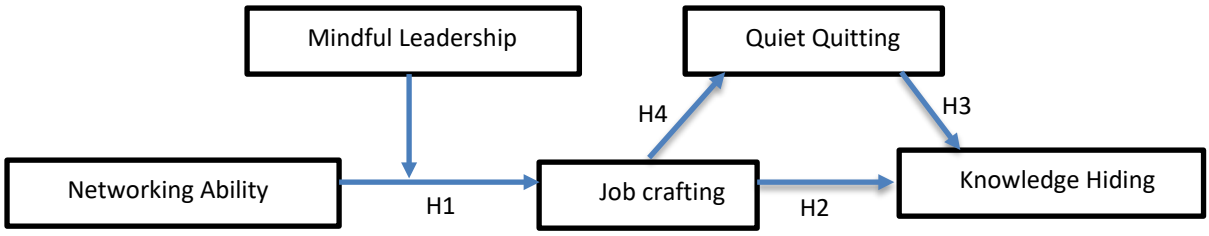


Figure 1 Research Framework

3 Research Methods

This study investigated full-time employees in Taiwan and collected data through a questionnaire survey, obtaining 321 valid samples. The mean age of the participants was 34.57 years (SD = 8.22), and the mean total years of work experience was 10.97 years (SD = 8.20). This study included gender, marital status, and education level as control variables.The mean values were 0.40, 0.37, and 1.24 (SD = 0.57).

Networking Ability was assessed using the Network Competence Scale from the Political Skill Inventory developed by Ferris et al. (2005). The Cronbach's α was 0.894. The Job crafting study used the Sora, Caballer and García (2018) scale, with 12 items, Cronbach's α = 0.893. Mindfulness leadership uses a modified set of items by Brown and Ryan (2003), consisting of 5 items, all on a Likert seven-point scale, where 1 represents "strongly disagree" and 7 represents "strongly agree", with Cronbach's α = 0.882. Quiet quitting was assessed using a scale modified from Galanis et al. (2023). Cronbach's α = 0.797. Knowledge hiding was assessed using the knowledge hiding scale developed by Connelly et al. (2012). The scale included three dimensions: feigning ignorance, avoidance hiding, and rationalization hiding. Cronbach's α = 0.957. All items were on a Likert seven-point scale, with 1 representing "strongly disagree" and 7 representing "strongly agree".

And used SPSS and Hayes' PROCESS for data analysis. PROCESS Model 7 was used to examine the relationships between networking skills, Job crafting, quiet quitting, and mindful leadership; PROCESS Model 4 was used to examine the relationships between Job crafting, quiet quitting, and knowledge hiding. All indirect effects were estimated using 5,000 bootstrap resampling.

4 Result

Descriptive statistics and correlation analysis showed that knowledge hiding was positively correlated with quiet quitting ($r = 0.260$, $p < 0.001$) and negatively correlated with job crafting ($r = -0.192$, $p < 0.001$). Networking ability was significantly positively correlated with job crafting ($r = 0.329$, $p < 0.01$) and significantly negatively correlated with quiet quitting ($r = -0.174$, $p < 0.01$). Job crafting was significantly negatively correlated with quiet quitting ($r = -0.457$, $p < 0.001$). Mindful leadership was significantly positively correlated with job crafting ($r = 0.172$, $p < 0.01$) and significantly negatively correlated with quiet quitting ($r = -0.198$, $p < 0.01$). Overall, the directions of the correlations were largely consistent with the theoretical expectations of this study.

Table 1 Correlation coefficients of variables

Variable	Mean	SD	1	2	3	4	5	6	7
1.Gender	--	--							
2. Marital Status	--	--	-0.016						
3. Education Level	--	--	0.085	-0.083					
4. Networking Ability	26.11	6.766	0.156**	0.097	0.074				
5. Job crafting	58.07	10.843	-0.052	0.143*	-0.005	0.329**			
6. Mindfulness leadership	24.16	6.687	-0.076	0.077	0.002	0.015	0.172**		
7. Quiet quitting	15.22	5.634	0.046	-0.115*	-0.078	-0.174**	-0.457**	-0.198**	
8. Knowledge hiding	27.50	13.718	0.083	0.004	0.048	0.008	-0.192***	0.162**	0.260***

Note：***p < .001, **p < .01, *p < .05。

This study used SPSS and Hayes' PROCESS macro for data analysis. PROCESS Model 7 was employed to examine the relationships among networking ability, job crafting, quiet quitting, and mindful leadership, whereas PROCESS Model 4 was used to test the relationships among job crafting, quiet quitting, and knowledge hiding. The results indicated that networking ability had a significant positive effect on job crafting ($B = 0.9493$, $SE = 0.2063$, $t = 4.6027$, $p < .001$), supporting H1. In addition, job crafting had a significant negative effect on quiet quitting ($B = -0.2269$, $SE = 0.0278$, $t = -8.1581$, $p < .001$), supporting H2. By contrast, the direct effect of

networking ability on quiet quitting was not significant (B = -0.0177, SE = 0.0449, t = -0.3950, p = .6931). Regarding the moderating effect, the interaction between networking ability and mindful leadership had a significant negative effect on job crafting (B = -0.0254, SE = 0.0114, t = -2.2396, p = .0258).

As shown in Table 2, networking ability had a significant positive effect on job crafting (B = 0.9493, SE = 0.2063, t = 4.6027, p < .001), supporting H1. Job crafting had a significant negative effect on quiet quitting (B = -0.2269, SE = 0.0278, t = -8.1581, p < .001), supporting H2. The direct effect of networking ability on quiet quitting was not significant (B = -0.0177, SE = 0.0449, t = -0.3950, p = .6931). In addition, the interaction between networking ability and mindful leadership had a significant negative effect on job crafting (B = -0.0254, SE = 0.0114, t = -2.2396, p = .0258), indicating that mindful leadership weakened rather than strengthened the positive relationship between networking ability and job crafting.

Table 2 Moderated Mediation Analysis Results (Model 7)

Path	coefficient (β)	SE	t	p value	95% CI
Networking Ability→Job crafting	0.9493	0.2063	4.6027	<0.001	[0.544, 1.355]
Networking Ability→Quiet quitting	-0.0177	0.0449	-0.395	0.693	[-0.106, 0.071]
Mindful leadership→Job crafting	0.4373	0.3154	1.3867	0. 1665	[-0.183, 1.058]
Networking Ability * Mindful leadership→ Job crafting	-0.0254	0.0114	-2.24	0.0258	[-0.048, -0.003]
Job crafting→Quiet quitting	-0.2269	0.0278	-8.158	<0.001	[-0.282, -0.172]

Note: Sex, marital status, and education were included as control variables. Unstandardized coefficients are reported.

Table 3. Mediation Analysis Results (PROCESS Model 4)

Path	coefficient (β)	SE	t	p value	95% CI
Job Crafting → Quiet Quitting	-0.2305	0.0262	-8.7900	< .001	[-0.282, -0.179]
Quiet Quitting → Knowledge Hiding	0.5515	0.1489	3.7036	.0003	[0.259, 0.845]
Job Crafting → Knowledge Hiding (direct effect)	-0.1182	0.0773	-1.5281	.1275	[-0.270, 0.034]
Job Crafting → Knowledge Hiding (total effect)	-0.2453	0.0707	-3.4702	.0006	[-0.384, -0.106]
Indirect effect via Quiet Quitting:	-0.1271	0.0480			[-0.226, -0.036]

Note. Sex, marital status, and education were included as control variables. Unstandardized coefficients are reported.

To further examine the relationship among job crafting, quiet quitting, and knowledge hiding, this study employed Hayes' PROCESS Model 4. The results showed that job crafting had a significant negative effect on quiet quitting (B = -0.2305, SE = 0.0262, t = -8.7900, p < .001), indicating that employees who engaged in higher levels of job crafting were less likely to exhibit quiet quitting tendencies. In turn, quiet quitting had a significant positive effect on knowledge hiding (B = 0.5515, SE = 0.1489, t = 3.7036, p = .0003), supporting H3.

The total effect of job crafting on knowledge hiding was significant (B = -0.2453, SE = 0.0707, t = -3.4702, p = .0006). However, after quiet quitting was included in the model, the direct effect of job crafting on knowledge hiding became non-significant (B = -0.1182, SE = 0.0773, t = -1.5281, p = .1275). Furthermore, the indirect effect of job crafting on knowledge hiding through quiet quitting was significant (Effect = -0.1271, BootSE = 0.0480, 95% BootCI = [-0.2265, -0.0368]). Because the bootstrap confidence interval did not include zero, the mediation effect was confirmed. These findings indicate that quiet quitting fully mediated the relationship between job crafting and knowledge hiding.

5 Discussion

The results of this study show that Networking Ability significantly promote Job crafting, which in turn significantly reduces quiet quitting. This aligns closely with Generation Z's work values; they value self-actualization, personal growth, fair compensation, and a positive organizational atmosphere. Furthermore, they believe that quiet quitting are unavoidable in this era (Chang & Chen, 2023). Networking ability significantly promote Job crafting, which in turn significantly reduces quiet turnover. This supports the COR theory's explanation of the proactive resource management process: employees with higher networking ability are more effective at obtaining support, information, and assistance, and translate these resources into proactive adjustments to work tasks, interactions, and cognition, thereby improving their resource status and reducing psychological detachment.

To enhance knowledge sharing within an organization, companies should emphasize coordination through interpersonal networks to create an environment conducive to knowledge sharing (Tung, 2004). When

employees leave quietly, the organization loses not only interpersonal communication but also the high-quality knowledge input supporting the operation of intelligent systems. This can result in aging knowledge bases, biased decision-making models, and decreased effectiveness of intelligent management tools. From this perspective, quiet departures in the AI era are not only a human resources issue but also a problem of intelligent knowledge governance.

This study found that mindful leadership has a significant negative moderating effect on the relationship between networking ability and Job crafting. Specifically, when the level of mindful leadership is lower, the promoting effect of networking ability on Job crafting is stronger. This result may suggest that, under less supportive contexts, employees with high networking ability are more likely to proactively adjust their work by relying on their own relationship resources; in contrast, under higher mindful leadership contexts, the support and security provided by supervisors may partially replace the necessity for employees to mobilize their personal network resources. When mindful leadership is low, the positive impact of networking skills on Job crafting is stronger; when mindful leadership is high, this effect is weaker. One possible explanation is that when situational support is low, employees with high networking skills need to rely more on their relationship resources to proactively adjust their work; conversely, when supervisors have provided high levels of support and psychological safety, the marginal utility of individual networking skills may decrease.

6 Conclusions

This study makes several theoretical contributions. First, compared to most studies that emphasize the consequences of quiet departure, this study further explores how organizations and employees can mitigate its occurrence through personal resources and proactive behaviors. It expands the application of COR theory, demonstrating that networking ability and job repositioning constitute a meaningful mechanism chain that helps employees prevent quiet departure. This shows that COR can explain not only defensive exits but also the proactive process of resource investment and gains. Second, this study incorporates knowledge hiding into the follow-up results of quiet quitting, expanding the explanatory power of quiet quitting research in the field of knowledge management. This indicates that quiet quitting are not only a low-input phenomenon, but also an important mechanism that weakens knowledge flow, knowledge contribution, and the transfer of tacit knowledge. Finally, this study links psychological withdrawal with the operation of intelligent organizations. The interruption of knowledge flow caused by employee psychological withdrawal in an organization may further affect the knowledge base and decision-making quality upon which intelligent systems rely, thus opening up new avenues for discussion on the application of COR theory in the context of intelligent management.

From a practical perspective, this study shows that organizations cannot rely solely on institutional controls or performance pressure to reduce quiet quitting and knowledge hiding; they should also emphasize employee access to resources and flexibility in job crafting. First, companies can enhance employees' networking and resource mobilization capabilities by encouraging cross-departmental interaction, building supportive networks, and promoting peer collaboration. Second, managers should provide more work autonomy and role adjustment opportunities, encouraging employees to reinvent their jobs to enhance their sense of control, purpose, and sustainable engagement.

Urthermore, organizations should view quiet quitting as a knowledge risk. The effectiveness of intelligent systems depends not only on the amount of data but also on the freshness, context, and willingness to continuously update the knowledge input. If employees are generally in a state of low engagement or knowledge silencing, even if the intelligent system appears to be operating continuously, the quality of its output, the accuracy of its predictions, and the credibility of its management recommendations may gradually decline.

This study, based on COR theory, the results show that networking capabilities significantly promote job reshaping, which in turn further reduces quiet departure; quiet departure significantly increases knowledge hiding and plays a fully mediating role between job reshaping and knowledge hiding. Mindful leadership has a marginally significant moderating effect on the relationship between networking capabilities and job reshaping.

The research results indicate that networking skills significantly promote job repositioning, which in turn significantly reduces the likelihood of employees leaving quietly. This confirms the resource conservation theory's explanation of proactive resource management. Overall, this study not only supplements the theoretical framework of the formation and subsequent impact of quiet quitting, but also reminds organizations to re-understand quiet quitting from the perspectives of resource management, knowledge flow, and intelligent management. By promoting job reshaping, strengthening interpersonal resource mobilization, and establishing intelligent knowledge governance mechanisms, organizations can reduce employee psychological withdrawal and mitigate its negative impact on knowledge management and intelligent management systems.

7 Ethics Declaration

This study collected data using anonymous questionnaires. Before completing the questionnaire, participants were informed of the research purpose, data usage, and confidentiality principles. This study did not collect any personally identifiable information; all data was used solely for academic research analysis, and the results were presented in a holistic statistical manner to protect participant privacy and research ethics.

8 AI Declaration

During the preparation of this manuscript, generative CHATGPT were used to assist with Chinese–English translation, and language polishing. These tools were not involved in data collection, statistical analysis, interpretation of results, or conclusion development. The research design, theoretical development, data analysis, interpretation of findings, and final approval of the manuscript were completed by the author, who assumes full responsibility for the content.

References

Connelly, C. E., Zweig, D., Webster, J., & Trougakos, J. P. (2012). Knowledge hiding in organizations. *Journal of Organizational Behavior*, *33*(1), 64–88.

Chang, C. J., & Chen, M. Y. C. (2023). Why generation Z is quiet quitting? A preliminary study of qualitative research. Journal of Technology and Human Resource Education, *9*(3), 1–32.

Gallup. (2023). *State of the global workplace: 2023 report*. Gallup Press.

Hamouche, S., Koritos, C., & Papastathopoulos, A. (2023). Quiet quitting: Relationship with other concepts and implications for tourism and hospitality. *International Journal of Contemporary Hospitality Management*, *35*(12), 4297–4312.

Ho, C. F., et al. (2022). Knowledge Management Strategies Based on Environment-Dependent and Organizational Learning Patterns. *Sun Yat-sen Management Review*, *30*(1), 203–240.

Hobfoll, S. E. (1989). Conservation of resources: A new attempt at conceptualizing stress. *American Psychologist*, *44*(3), 513–524.

Hobfoll, S. E. (2011). Conservation of resources theory: Its implication for stress, health, and resilience. In S. Folkman (Ed.), *The Oxford handbook of stress, health, and coping* (pp. 127–147). Oxford University Press.

Hur, W.-M., Shin, Y., & Kim, J. Y. (2024). Service employees' mindfulness and job crafting amid COVID-19: The roles of resilience, organizational health climate, and health-oriented leadership. *Current Psychology*, *43*(18), 16979–16991.

Kachhap, V., & Singh, T. (2024). Quiet quitting: A comprehensive exploration of hidden problems. *Development and Learning in Organizations: An International Journal*, *38*(5), 23–26.

Kim, K. T., & Sohn, Y. W. (2024). The impact of quiet quitting on turnover intentions in the era of digital transformation: The mediating roles of job satisfaction and affective commitment, and the moderating role of psychological safety. *Systems*, *12*(11), Article 460.

Klotz, A. C., & Bolino, M. C. (2023, October 3). When quiet quitting is worse than the real thing. *Harvard Business Review*.

Kumar, M., Carvalho, S., & Carvalho, C. (2022). Employee networking behavior: Sources, challenges, and support. *Advances in Developing Human Resources*, *24*(2), 142–150.

Lu, C. Y., & Chen, L. Y. (2020). The Impact of Job Crafting on Work Engagement and Job Performance: The Moderated Mediation of Feedback Seeking Behavior. *Journal of Business Administration*, *45*(1), 91–119.

Liu, H. R., & Chen, H. Y. (2026). Concept Analysis of Quiet Quitting. *Hu Li Za Zhi*, *73*(2), e26211.

Liu, Y., et al. (2020). Internal selling: Antecedents and the importance of networking ability in converting internal selling behavior into salesperson performance. *Journal of Business Research*, *117*, 176–188.

Patel, P. C., et al. (2025). A multidimensional quiet quitting scale: Development and test of a measure of quiet quitting. *PLOS ONE*, *20*(4), e0317624.

Scheyett, A. (2022). Quiet Quitting. *Soc Work*, *68*(1), 5–7.

Shen, Y., et al. (2025). A meta-analysis of knowledge hiding behavior in organizations: Antecedents, consequences, and boundary conditions. *Journal of Business Research*, *186*, Article 114963.

Song, C. N., Lin, H. C., & Wang, S. N. (2023). Top Leader-Top Management Team Exchange and Management Innovation: Moderated-Mediation Effect of Leader-Member Exchange Differentiation. *Journal of Human Resource Management*, *23*(1), 27–56.

Sutcliffe, K. M., Vogus, T. J., & Dane, E. (2016). Mindfulness in organizations: A cross-level review. *Annual Review of Organizational Psychology and Organizational Behavior*, *3*, 55–81.

Tan, N., Peters, E. K., & Reb, J. (2023). Effects of a mindfulness-based leadership training on leadership behaviors and effectiveness. *Mindfulness*, *14*(9), 2181–2194.

Tims, M., & Bakker, A. B. (2010). Job crafting: Towards a new model of individual job redesign. *SA Journal of Industrial Psychology*, *36*(2), Article 841.

Tung, Y. C. (2004). Antecedents of Knowledge Sharing Willingness: The Moderating Effect of the Threat Variable. *Journal of Human Resource Management*, *4*(3), 117–137.

Vonderlin, R., et al. (2021). Effectiveness of a mindfulness- and skill-based health-promoting leadership intervention on supervisor and employee levels: A quasi-experimental multisite field study. *Journal of Occupational Health Psychology*, *26*(6), 613–628.

Wang, Z., Shi, L., & Wang, L. (2021). Does leader mindfulness influence voice behavior? Leader–member exchange as a mediator. *Social Behavior and Personality: An International Journal*, *49*(8), 1–8.

Wei, T. Y., Jaw, B. S., & Hsiao, Y. C. (2022). Work Values, Job Insecurity, Job Pleasure, and Job Involvement of Employees influencing Intention to Stay in Startup company. *International Journal of Quantitative Analysis and Research*(1), 32–48.

Weng, L. C., & Li, Y. C. (2015). A Multi-level Investigation on the Relationships between Transformational Leadership and Knowledge Sharing: The Social Dilemma Perspective. *International Journal of Commerce and Strategy*, *7*(1), 53–70.

Wickramaratne, W. (2020). Role of employee networking behavior in predicting the relationship between perceived supervisor support and career satisfaction. *Peradeniya Management Review*, *2*(2), 79–102.

Wrzesniewski, A., & Dutton, J. E. (2001). Crafting a job: Revisioning employees as active crafters of their work. *Academy of Management Review*, *26*(2), 179–201.

Yiu, T. F., Huang, Y. C., & Wang, K. Y. (2025). The "Quiet Quitting" in Nursing Workplaces and Coping Strategies. *Hu Li Za Zhi*, *72*(3), 95–101.

Zhang, F., & Parker, S. K. (2019). Reorienting job crafting research: A hierarchical structure of job crafting concepts and an integrative review. *Journal of Organizational Behavior*, *40*(2), 126–146.

Knowledge Dissemination Strategies for Circular Economy: Evidence from Italian Local Governments

Miriana Ferrara[1], Nicola Capolupo[2]
[1] University of Salerno, Fisciano, Italy
[2] San Raffaele Roma University, Rome, Italy

Abstract

The circular economy (CE) has emerged as a central paradigm in the governance of urban sustainability, assigning local governments a crucial role in the transition towards regenerative models of production and consumption. However, the success of CE initiatives depends not only on the design of regulatory strategies and instruments, but also on the ability of public institutions to disseminate knowledge effectively to citizens and stakeholders. However, the literature has devoted greater attention to the design of CE policies rather than to the processes through which such policies are made accessible and socially embedded.

This study investigates the knowledge dissemination strategies adopted by three Italian regional capitals (Florence, Turin and Genoa) through their institutional channels. The analysis draws on the framework of Green et al. (2009), which articulates the dissemination process into five constituent dimensions: the source, the content, the medium, the user and the context. The research design is based on a multiple-case approach, involving the collection of qualitative secondary data, primarily institutional websites, project documentation and official reports, analysed through direct qualitative content analysis.

The findings reveal that, while sharing a formal commitment to the circular transition as signatories of the European Circular Cities Declaration, the three cities are developing structurally distinct dissemination strategies. Florence adopts a civic dissemination strategy towards citizens' behavioural compliance. Turin develops a business-oriented ecosystemic strategy, in which knowledge is co-produced within a permanent urban innovation platform. Genoa presents the most complex architecture, characterized by a dual-track system that explicitly differentiates the registers intended for citizens and economic actors.

The study contributes to the literature by extending the application of Green et al's framework to the domain of CE governance in local public administration, demonstrating how choices across the five dimensions do not operate independently but constitute coherent epistemic architectures that encode distinct governance philosophies.

Keywords: Circular Economy, Circular Cities, Public Administrations, knowledge dissemination

Paper type: Academic research paper

1 Introduction

Circular economy (CE) has emerged as one of the defining paradigms of contemporary sustainability governance, offering a systemic alternative to the linear "take-make-dispose" model that has underpinned industrial production for decades (MacArthur, 2013; Geissdoerfer et al., 2017). By designing out waste, maintaining materials and products in circulation at their highest utility, and regenerating natural systems, CE provides an integrated framework through which economic growth can be progressively decoupled from resource consumption and environmental degradation (Kirchherr et al., 2017). In this evolving policy landscape, cities and metropolitan areas have assumed a central governance role: urban territories concentrate both the resource flows that CE seeks to reconfigure (waste, mobility, construction, consumption) and the institutional density required to translate strategic commitments into operational practice (Fratini et al., 2019; Paiho et al., 2020).

Local governments are accordingly positioned as key actors in CE governance, tasked not only with designing regulatory instruments and investment strategies but also with managing the communicative and participatory dimensions of the transition. CE initiatives, particularly those addressing urban waste and sustainable consumption, cannot succeed without the active involvement of citizens as both behavioural actors and co-producers of sustainable practices (Moreau et al., 2017). This civic dimension of CE governance places knowledge dissemination - the capacity of public institutions to share actionable, meaningful knowledge with diverse stakeholder communities - at the heart of any effective urban CE strategy. Yet this dimension remains comparatively underexplored in the scholarly literature, which has devoted substantially greater attention to the design of CE strategies than to the communicative and epistemic processes through which these strategies are made accessible and socially embedded (Pozzetto & Leoni, 2025; Ducci, 2025).

The present study addresses this gap by investigating how three Italian regional capitals - Turin, Florence, and Genoa - disseminate CE knowledge to citizens and other stakeholders through their institutional websites and associated communication channels. To analyze knowledge dissemination strategies in a theoretically systematic way, this study draws on the knowledge communication framework developed by Green et al. (2009). The framework identifies five constitutive and interrelated elements of any dissemination process: the source, who produces the knowledge; the content, what knowledge is shared and how it is framed; the medium, through which channels knowledge is conveyed; the user, who is expected to receive and apply the knowledge; and the context, the institutional and community environment within which dissemination occurs. Applied to the institutional websites of local governments, this framework enables a structured comparative analysis that moves beyond descriptive accounts of CE communication towards an examination of the underlying strategic logics shaping how cities conceptualize their roles as knowledge disseminators.

The paper contributes to the literature in two principal ways. Theoretically, it extends the application of the framework to the domain of CE governance in local public administration, demonstrating its analytical purchase as a tool for mapping knowledge dissemination strategies across diverse institutional contexts. Empirically, it provides comparative evidence on how Italian municipalities with comparable CE commitments nonetheless develop markedly distinct approaches to civic knowledge sharing. The paper is organized as follows. Section 2 reviews the relevant literature and presents the theoretical framework. Section 3 describes the research methodology. Section 4 presents the within-case and cross-case analysis. Section 5 discusses findings and implications, and Section 6 offers brief concluding remarks.

2 Literature Review and Theoretical Background

2.1 The role of knowledge dissemination for circular economy

Research on sustainability transitions highlights that the successful implementation of CE strategies depends not only on policy design and material infrastructures, but also on how knowledge is actively disseminated and internalized by diverse actors within a territory (Bourdin & Jacquet, 2025; Kasmi et al., 2022). Circularity involves complex changes in behaviors, organizational routines, and market practices (Klein et al., 2022); thus, knowledge dissemination is a core governance function that shapes how stakeholders interpret, prioritize, and operationalize CE principles in practice (Mele et al., 2025).

Dissemination in this context is not a unidirectional information broadcast but a deliberate process of engagement, translation, and sense-making. It involves tailoring information to the specific knowledge levels, expectations, and motivational frames of different actor groups, including citizens, firms, and civic organisations (Avelino et al., 2016). Research on sustainability communication shows that generic informational campaigns rarely translate into behavioural change when they fail to align with individuals' pre-existing values, ideological orientations, and perceived behavioural control, which are particularly influential in circular economy practices such as recycling, reuse, and sharing (Shah & Yang, 2025, Lorincová & Osvaldovà, 2025).

In municipal governance, the capacity to disseminate circular-economy knowledge intersects with core administrative functions such as coordination, openness, and institutional trust, which in turn influence citizens' willingness to participate in sustainability actions. Moreover, interactive dissemination practices, those that go beyond one-way information provision and create dialogue spaces, can facilitate collective and reflective learning among stakeholders, helping to co-develop shared understandings, align expectations, and adapt policies responsively (Jacobi & De Souza, 2021). Such collaborative learning processes are increasingly recognized as vital in environmental governance, where everyday actions by households and local businesses shape the overall effectiveness of policy interventions

A further theme concerns the digitalization of knowledge dissemination. Digital platforms, open data initiatives, and social media are increasingly deployed by local governments not just to distribute information but to create spaces for dialogue, feedback, and collaborative problem-solving. These tools can reduce information asymmetries and enhance accessibility to CE knowledge (Hossain et al., 2016; Ranta et al., 2018), although their effectiveness depends on design choices that acknowledge digital inclusiveness and trust dynamics (Cerić & Ivić Jazvec, 2026). Hence, knowledge dissemination for CE should be analyzed as a governance mechanism that influences cognitive frames, social norms, and institutional relationships, rather than as a peripheral communication task. Assessing how municipalities craft, tailor, and mediate CE-related knowledge flows therefore becomes essential to understanding CE implementation processes in urban contexts.

2.2 A knowledge dissemination framework

To provide a structured analytical lens for examining CE knowledge dissemination in local government contexts, this study adopts the framework developed by Green et al. (2009). Originally conceived in the field of public health as a model for understanding the conditions under which research-based knowledge is diffused, utilized, and integrated into practice, the framework has since been taken up in management and organizational studies as a generalized interpretive schema for analyzing knowledge transformation processes (Serrat, 2017). Its relevance to the present study rests on its explicit conceptualization of dissemination as a multi-dimensional and relational process rather than a simple act of information transfer, and its consequent capacity to illuminate the strategic choices embedded in any act of institutional knowledge communication.

The framework identifies five constitutive elements of the dissemination process, each of which shapes the conditions under which knowledge is effectively produced, transmitted, received, and applied. To better describe and contextualize them, this study draws on the definitional elaborations provided by the National Center for the Dissemination of Disability Research (NCDDR, 1996), which operationalizes each dimension in terms of its practical functions within a dissemination process.

The first element is the source, which refers to the actor or institution originating the knowledge. The credibility, authority, and perceived expertise of the source significantly influence whether disseminated knowledge is perceived as trustworthy by its intended recipients; in public sector contexts, the institutional position of the source, whether a municipality acting alone, a public-private partnership, or a multi-stakeholder network, shapes the epistemic authority with which knowledge is invested.

The second element is the content, which encompasses the nature, scope, framing, and cognitive objectives of the knowledge being shared. Content may range from prescriptive and behavioral-directive to exploratory and project-based, inviting businesses and institutions to co-construct novel circular solutions. The third element is the medium, which refers to the channels and formats through which knowledge is conveyed. The choice of medium affects not only the reach of dissemination but also its depth, interactivity, and the degree to which it can accommodate diverse knowledge needs. Institutional websites, social media, physical spaces, and documentary artifacts each carry distinct affordances and limitations as knowledge dissemination tools (Pozzetto & Leoni, 2025).

The fourth element is the user, which refers to the intended or actual recipients of the disseminated knowledge, including their prior knowledge, motivations, and contextual capacity to absorb and act upon it. The conceptualization of the user reflects deeper assumptions about the role of citizens and other stakeholders in CE governance: whether they are positioned as passive recipients of prescriptive guidance, as active participants in co-creation processes, or as economic actors whose engagement is sought based on professional or entrepreneurial interest.

The fifth and final element is the context, which encompasses the institutional, social, and political environment within which dissemination occurs. Contextual factors such as governance structures, formal policy commitments, civic culture, and administrative capacity all condition the production, circulation, and reception of CE knowledge, and must be considered in any comparative analysis of dissemination strategies.

These five dimensions constitute an integrated analytical framework that is particularly well-suited for the examination of institutional communication channels as artifacts of CE public knowledge dissemination (Fig.1).

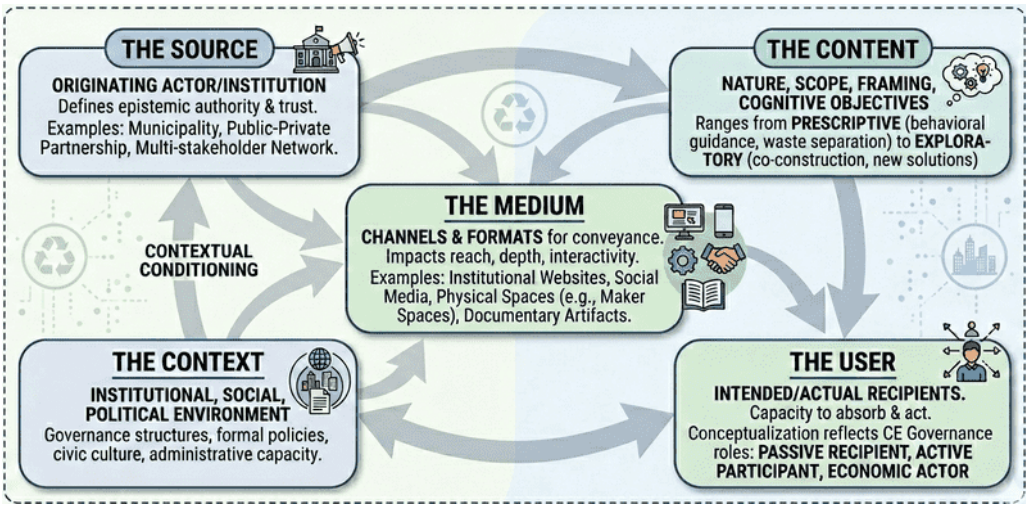


Figure 1: Circular city knowledge dissemination framework, authors 'elaboration through Nano Banana 2 AI software

3 Research Methodology

3.1 Research design and research context

Since the objective of the study is to understand knowledge dissemination strategies developed by diverse local governments to share their CE initiatives, the present study adopts a multiple case study approach (Yin, 2009). Indeed, this method is appropriate for collecting evidence about contemporary and complex real-life that requires deeper investigation by comparing and pointing out similarities and differences of the case studies (Yin, 2018). In this logic, the study first unfolds the knowledge dissemination elements according to the framework provided by Green et al. (2009), then conducts a cross-case analysis to uncover different strategic approaches and identify the main knowledge dissemination strategy developed by local governments.

The cases were selected according to a non-probability sampling (i.e., purposive) (Rapley, 2014), that matched the following parameters:

- being from the same country, to unify national legislation and socio-cultural backgrounds, then implementing the replication logic identifying similarities and difference;
- being from Italy, which ranks as a leading country among the largest EU economies for CE initiatives;
- being regional capital, to have comparable level of organizational complexity;
- being signatories to the European Circular Cities Declaration, since this shows the formal commitment that these cities have taken towards CE initiatives;
- being limited to three cases in line with the replication logic of multiple case study design, which recommends a manageable number of cases for in-depth cross-case comparison rather than breadth of coverage (Yin, 2018).

Three Italian local governments were selected for the analysis, i.e., Florence, Turin and Genoa.

Florence has institutionalized its efforts towards CE through the Florence Circular City Strategy 2021-2023, which has seen the development of two main projects named Florence Plastic Free and Florence Circular City.

Turin has seen its commitment to CE initiatives as part of a strategic approach to innovating the city. In this regard, the local government has promoted Turin City Lab, which is an urban digital platform, established to create a streamlined environment in which businesses, start-ups, research bodies and public sector organizations can test innovative solutions in real-world conditions.

Genoa has developed and implemented a participatory CE strategy as part of the C-City project. The strategy considers the short, medium and long-term perspective, and on the upstream objective of reducing raw material and produce usage in the city, aiming long-term towards a zero-waste city.

3.2 Data collection and data analysis

Secondary qualitative data were collected, as they were produced by institutional actors for other purposes than those of this research and are publicly accessible without intermediaries. No primary data collection was carried out, in line with the study's focus on the dissemination of produced knowledge, rather than on individual perceptions or behaviors. They were collected in January 2026 mainly through local governments official websites, since these are considered as primary public channels for the dissemination of knowledge, representing institutional artifacts through which administrations seek to make information on circular economy policies and initiatives understandable and usable (Pozzetto & Leoni, 2025; Ducci, 2025). Beyond the screening of official websites (torinocitylab.it/; firenzecittacircolare.it; c-citygenovacircolare.eu), other analyzed sources included: the most updated Circular Cities Declaration Report (which refers to 2024), websites devoted to individual CE initiatives, the documentation for each CE initiative (e.g., Florence Plastic Free Action Plan, Genoa 2050 Action Plan, Turin Urban Innovators Guide).

The data were analyzed using qualitative directed content analysis (Hsieh & Shannon, 2005). This approach was chosen because it allows existing theory or theoretical frameworks to guide the initial coding process, providing a well-defined set of analytical dimensions that are then applied to the empirical context. This approach is particularly appropriate for the present study, in which the framework provided by Green et al. (2009) defines the main categories that are then conceptualized in a novel context. Therefore, the coding scheme was structured around the five dimensions identified by Green et al. (2009): source, content, medium, user and context. The analysis was conducted in two sequential stages. Firstly, a within case analysis was carried out for each city separately, examining how each of the five dimensions is represented. Secondly, a cross-case analysis was conducted, identifying systematic similarities and differences, finally extrapolating the main knowledge dissemination strategy for each local government. To enhance the reliability of the coding process,

the two authors independently coded a subset of the data and subsequently compared and reconciled their interpretations through iterative discussion until consensus was reached on all categorizations.

4 Results

4.1 Within case analysis

4.1.1 Florence

As aforementioned, Florence CE initiatives are embedded in two main activities: Florence Circular City and Florence Plastic Free.

Regarding Florence Circular City, the knowledge source is represented by the Municipality of Florence and Alia Servizi Ambientali S.p.A., the private operator responsible for environmental services within the city. The partnership model positions knowledge production as a shared responsibility between these two actors. The content of knowledge, disseminated primarily through dedicated websites, mainly regards the description of the city's circular economy activities in the waste sector, the provision of detailed guidance on how to carry out separate waste collection, and the indication of the collection points distributed across the urban area. The medium, beyond the institutional website, extends beyond the website and includes social media channels (Instagram and Facebook), radio broadcasts, face-to-face community meetings (coordinated by the A-team) and a dedicated mobile app, reflecting a deliberate multi-platform strategy aimed at maximising the reach of citizens. The target user is primarily the individual citizen in their capacity as an urban resident and waste producer.

Florence Plastic Free is a second initiative launched in partnership with the waste collection service provider and the water service provider. Beyond these actors, the initiative further involves knowledgeable partners, such as trade associations in the tourism sector, operators of food and non-food markets, trade associations for commercial activities, and managers of public and private canteens, recognising the systemic role of economic actors in reducing plastic. The content is structured around four explicitly stated objectives: reducing the consumption of single-use plastic, promoting good practices through information and incentives, the proper management of plastic waste, and the cleaning of urban areas. The medium is related to a dedicated page on the municipal environmental services website (ambiente.comune.fi.it), which hosts the initiative's action plan. Other forms of knowledge dissemination regard the promotion of targeted awareness campaigns in schools and public offices, and the organisation of waste collection days in sensitive urban areas in collaboration with schools and local associations. The main users include primarily individual citizens.

The context is characterized by formal commitment to CE initiatives, through the Florence Circular City Strategy, however it lacks a unified strategic document.

4.1.2 Turin

The source of knowledge production is the Local Department for Ecological and Digital Transition, working in synergy with Torino City Lab, bringing together public administration, private enterprises and academic stakeholders in real-world experimentation. Therefore, the knowledge source emerges to be produced by not a single institutional actor but by a multi-stakeholder network that co-produces and co-disseminates knowledge in a systematic and formalised manner.

The content of knowledge regards focusing on the development of circular urban supply chains, local infrastructure for reuse and repair, and products and services based on circular design. The aim of the content is not to teach but to engage partners to co-produce CE initiatives with, then shaping a collaborative and project-based approach.

The medium used are the institutional website and social network pages which function primarily as vehicles for calls to action aimed at potential business partners. These channels are not designed to disseminate knowledge or best practices to the public; on the contrary, they assume a proactive audience capable of responding concretely to institutional push. Consequently, the target audience consists almost exclusively of players from the business world such as companies, start-ups and research bodies, while individual citizens are largely absent as recipients of communication on the circular economy.

The institutional context is characterized by the institutional effort of the local government to shape an innovative city, capable of integrating technological experimentation, social impact and ecological transition.

4.1.3 Genoa

In terms of the source, knowledge production is anchored to the Municipality of Genoa as the central institutional actor. The distinctive feature of the Genoese model, however, lies in the participatory and multi-

stakeholder nature of the knowledge production process: the city's CE strategy was developed through a process co-designed with citizens, the third sector, businesses and major economic operators, creating a source that is not merely institutional but ecosystemic and deliberately inclusive (ICLEI, 2024).

The content disseminated through Genoa's institutional channels is organised into seven thematic domains (i.e., energy, food, the sea, textiles, construction, mobility, materials and products) reflecting a specific effort in each sector. The content to convey practical and replicable knowledge includes best practices, neighbourhood-scale intervention models, and tools for monitoring and governing the circular transition.

In terms of its medium, Genoa is building a communication infrastructure that stands out for its combination of physical and digital spaces for knowledge exchange. The Circular Hub, located in the city centre, is an on-site and digital space of knowledge exchange among different stakeholders to co-create urban circular economy microprojects. The Circular Desk represents a second physical tool for knowledge diffusion, specifically dedicated to businesses: it functions as a technical facilitation desk that builds relationships between actors for CE initiatives and provides a preliminary assessment for open innovation activities linked to circular processes. The communication toolkit is complemented by the dedicated institutional website (c-citygenova.eu), editorial articles and structured events, including the Circular City Forum 2026, which Genoa will host as the host city.

The user dimension is explicitly dual and hierarchically differentiated. Citizens are engaged through the Circular Hub and its digital platform, which involves them not as passive recipients of prescriptive content but as active participants in the development of urban-scale circular micro-projects. Businesses, on the other hand, are directed towards the Circular Desk, which supports them on their path towards circularity and enables them to propose projects and share best practices.

The institutional context moves around the Genoa 2050 Action Plan, approved by municipal resolution, which provides the CE strategy with a stable and verifiable governance framework.

4.2 Cross-case analysis

The cross-case comparison reveals that, despite sharing a formal commitment to circular economy transition as signatories of the European Circular Cities Declaration, the three cities exhibit markedly distinct characteristics across all five dimensions framework. These differences track diverse knowledge dissemination strategies for each local government (Tab.1).

With regard to the nature of the source, all three cities involve the local municipality as the central institutional actor but differ significantly in the structure of knowledge production. Florence operates through a public-private partnership model, in which knowledge is co-produced with private service providers. Turin has established a networked institutional production in which the administration operates within a permanent platform that integrates businesses, start-ups and research bodies. Genoa developed a multi-stakeholder ecosystemic model for which the municipality plays a key role in co-producing knowledge with both business actors and citizens.

In terms of content, the clearest difference emerges between Florence on the one hand, and Genoa and Turin on the other. Florence produces prescriptive-behavioral knowledge, focused on individual daily actions. Turin develops content of a project-based and collaborative nature, in which knowledge is not transmitted but co-constructed with economic actors. Genoa adopts a strategic-programmatic approach, applying the CE across seven thematic domains within a deliberated action plan. This variation in the type of content signals profound differences in the very conception of dissemination: informing, planning or co-constructing.

Regarding the medium, Florence stands out for its greater diversification of channels, combining digital, physical and traditional tools in a multi-channel strategy aimed at maximum reach. Turin, by contrast, employs a minimal communication repertoire used mostly to convey entrepreneurial calls to action. Genoa builds a hybrid communication infrastructure, which combines a thematic platform dedicated to physical tools.

The user dimension represents the most marked divergence between the three cases and the one with the most direct implications for civic engagement. Florence directs its communication primarily with the individual citizen. Genoa explicitly adopts a dual model, establishing distinct communication channels for citizens and businesses. Turin effectively focuses on economic and institutional actors.

In terms of context, all three cities share the same formal starting point (i.e., adherence to the Circular Cities Declaration) but differ to the degree to which their commitment has been institutionalized. Florence maintains a general enabling framework, without a unified policy document to structure initiatives into a coherent strategy. Turin has built a permanent innovation infrastructure, that is Turin City Lab. Genoa has translated its formal commitment into a deliberate governance tool which provides EC initiatives with a stable and verifiable institutional foundation.

Florence adopts a diffusive civic dissemination strategy, in which source authority, behavioral content, multi-channel media, and a citizen-as-target user conception combine into a coherent logic of compliance-oriented civic mobilization. Turin has devised a business-oriented ecosystemic dissemination strategy, where a multi-stakeholder source, strategic-programmatic content, and a permanent institutional platform converge into a model in which knowledge is not transmitted but co-constructed. Genoa presents the most structurally complex strategy, characterized by a dual-track knowledge architecture that deliberately separates epistemic registers, directing action-oriented content to citizens while engaging businesses as strategic partners in CE governance. These three strategies do not constitute a hierarchy of effectiveness; rather, they reflect diverse yet coherent ways in which local governments translate a shared formal commitment to CE into structurally distinct knowledge governance arrangements.

Table 1: Results summary

Dimension	Florence	Turin	Genoa	
Source	Public-private partnership	Networked institutional production	Multi-stakeholder ecosystem	
Content	Prescriptive-behavioural	Project-based-collaborative	Strategic-programmatic	
Medium	Diffusive multi-channel	Action-oriented digital	Physical and Digital Platform	
User	Citizen	Businesses	Citizens and businesses	
Context	Formal commitment, no unified strategic document	Permanent innovation infrastructure (Torino City Lab)	Formal commitment + Action Plan (Genoa 2050)	

5 Discussion, Implications and Conclusions

The application of the Green et al. (2009) framework reveals that the five dimensions do not operate as independent conditions but as an integrated system in which choices along one dimension constrain and shape the others. What emerges is not simply variation in communication style, but three coherent epistemic architectures encoding distinct governance philosophies about who holds legitimate knowledge, for whom it is produced, and to what end.

Source and content are the dimensions most directly shaped by how cities conceptualize CE knowledge production. The shift from a public-private partnership source (Florence) to a multi-stakeholder ecosystemic model (Genoa) corresponds to a transition from institutional transfer to distributed co-production (Nonaka & Takeuchi, 1995). Correspondingly, the content variation, from prescriptive-behavioral to strategic-programmatic, maps onto a well-established typology in science communication between the "deficit model," which treats publics as lacking knowledge to be filled, and dialogic models that treat knowledge production as inherently participatory (Wynne, 1992; Felt et al., 2016).

Medium and users carry the most direct implications for civic engagement. The choice of channel presupposes and constructs an imagined audience: Florence serves a citizenry conceived as compliance targets; Turin serves professional co-innovators; Genoa differentiates explicitly between citizens and businesses. Mapping these onto Fung's (2006) typology, Florence approximates "managed participation," Turin resembles "stakeholder representation," and Genoa approaches "deliberative co-governance." The user dimension is the most politically significant of the five because it encodes the relationship between the local state and its publics, whether citizens are positioned as recipients or co-producers of knowledge (Bovaird, 2007).

Context conditions the coherence and durability of dissemination strategies over time. All three cities share the same formal commitment through the Circular Cities Declaration, yet this produces markedly different contextual architectures. Research on institutional capacity indicates that formal commitments generate sustained governance outcomes only when translated into binding programmatic instruments and dedicated organizational structures (Termeer et al., 2017). Genoa's Action Plan and Turin's permanent platform represent such translations; Florence's absence of a unified strategic document represents a contextual limitation that may limit the long-term scalability of its dissemination efforts.

This study makes two principal contributions to the emerging literature on knowledge governance for circular economy in urban contexts. Theoretically, it extends the application of the Green et al. (2009) knowledge dissemination framework beyond its original public health domain, demonstrating its analytical productivity as a lens for mapping the strategic logics embedded in institutional CE communication in local public administration. Empirically, it provides comparative evidence on how three Italian regional capitals (Florence, Turin, and Genoa) develop structurally distinct knowledge dissemination strategies despite sharing a common formal commitment to CE transition, yielding three coherent archetypes (diffusive civic, business-oriented ecosystemic, and dual-track knowledge architecture).

Several limitations must be acknowledged. The study relies exclusively on secondary, publicly available institutional sources, which by design reflect governance narratives rather than the actual reception, uptake, or practical effectiveness of disseminated knowledge. The purposive sample of three Italian cities, while analytically justified, limits the transferability of findings to other national governance contexts. These limitations outline a new research agenda. Future studies should integrate primary data collection, including citizen surveys, stakeholder interviews, and digital engagement analytics, to evaluate whether and how disseminated knowledge is received, interpreted, and translated into action. Comparative research extending beyond Italy and encompassing cities from different governance traditions and CE maturity levels, would enrich understanding of the structural and cultural determinants of dissemination strategy choice. Longitudinal designs would permit investigation of how strategies co-evolve with governance capacity and civic engagement over time. Finally, future research could productively examine the relationship between dissemination strategy type and measurable CE outcomes providing an empirical foundation for normative claims about the relative effectiveness of different knowledge governance approaches.

6 Ethics Declaration

Ethical approval was not required for this study.

7 AI Declaration

Nano Banana 2 AI software was used to generate Figure 1.

References

Avelino, F., Grin, J., Pel, B., & Jhagroe, S. (2016). The politics of sustainability transitions. Journal of Environmental Policy & Planning, 18(5), 557–567.

Bourdin, S., & Jacquet, N. (2025). The contribution of European cities and regions to the circular economy: An analysis of territorial strategies. Regional Studies, 59(1), 2563888.

Bovaird, T. (2007). Beyond engagement and participation: User and community coproduction of public services. Public administration review, 67(5), 846-860.

Cerić, A., & Ivić Jazvec, I. (2026). The role of blockchain technology in reducing information asymmetry and enhancing trust in circular construction. Sustainability, 18(5), 2328.

Circular Economy Network. (2025). Rapporto sull'economia circolare in Italia 2025. https://circulareconomynetwork.it/wp-content/uploads/2025/05/Rapporto-sulleconomia-circolare-in-Italia-2025.pdf

Ducci, G. (2025). Per una comunicazione pubblica (istituzionale) della sostenibilità: definizione e tipologia. Salute e Società, 24(2, Suppl.), 36–52.

Felt, U., Fouché, R., Miller, C. A., & Smith-Doerr, L. (Eds.). (2016). The handbook of science and technology studies. Mit Press.

Florence Plastic Free Action Plan. (n.d.). https://ambiente.comune.fi.it/firenze-plastic-free

Fratini, C. F., Georg, S., & Jørgensen, M. S. (2019). Exploring circular economy imaginaries in European cities: A research agenda for the governance of urban sustainability transitions. Journal of Cleaner Production, 228, 974–989.

Fung, A. (2006). Varieties of participation in complex governance. Public administration review, 66, 66-75.

Geissdoerfer, M., Savaget, P., Bocken, N. M. P., & Hultink, E. J. (2017). The circular economy–A new sustainability paradigm? Journal of Cleaner Production, 143, 757–768.

Genoa Action Plan. (2021). https://circularcitiesdeclaration.eu/fileadmin/user_upload/Images/Pages_Images/Genoa/GENOVA_2050_-_ActionPlan_light_Eng_town_approval_06052021.pdf

Green, L. W., Ottoson, J. M., Garcia, C., & Hiatt, R. A. (2009). Diffusion theory and knowledge dissemination, utilization, and integration in public health. Annual Review of Public Health, 30(1), 151–174.

Hossain, M. A., Dwivedi, Y. K., & Rana, N. P. (2016). State-of-the-art in open data research: Insights from existing literature and a research agenda. Journal of Organizational Computing and Electronic Commerce, 26(1–2), 14–40.

Hsieh, H. F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. Qualitative Health Research, 15(9), 1277–1288.

ICLEI. (2024). Circular Economy Declaration Report. https://circularcitiesdeclaration.eu/fileadmin/user_upload/Resources/CCD-Report-2024-FINAL-spreads.pdf

Jacobi, P. R., & De Souza, D. T. (2021). Environmental governance and social learning: Towards more reflexive interactions. Desenvolvimento e Meio Ambiente, 58, 932–946.

Kasmi, F., Osorio, F., Dupont, L., Marche, B., & Camargo, M. (2022). Innovation spaces as drivers of eco-innovations supporting the circular economy: A systematic literature review. Journal of Innovation Economics & Management, 39(3), 173–214.

Kirchherr, J., Reike, D., & Hekkert, M. (2017). Conceptualizing the circular economy: An analysis of 114 definitions. Resources, Conservation and Recycling, 127, 221–232.

Klein, N., Ramos, T. B., & Deutz, P. (2022). Factors and strategies for circularity implementation in the public sector: An organisational change management approach for sustainability. Corporate Social Responsibility and Environmental Management, 29(3), 509–523.

Lorincová, S., & Osvaldová, M. (2025). Pro-environmental behavior and attitudes towards recycling in Slovak Republic. Recycling, 10(4), 159.

MacArthur, E. (2013). Towards the circular economy. Journal of Industrial Ecology, 2(1), 23–44.

Mele, G., Del Vecchio, P., & Passiante, G. (2025). Platform-based ecosystems and knowledge management for the circular economy: An in-depth analysis of Italian cases. Journal of Knowledge Management, 29(9), 3011–3030.

Moreau, V., Sahakian, M., Van Griethuysen, P., & Vuille, F. (2017). Coming full circle: Why social and institutional dimensions matter for the circular economy. Journal of Industrial Ecology, 21(3), 497–506.

NATIONAL CENTER FOR THE DISSEMINATION OF DISABILITY RESEARCH (NCDDR). 1996. A Review of the Literature on Dissemination and Knowledge Utilization: purpose and overview. July 1996.

Nonaka, I. and Takeuchi, H. (1995) The Knowledge Creating Company: How Japanese Companies Create the Dynamics of Innovation. Oxford University Press, New York. https://doi.org/10.1016/0024-6301(96)81509-3

Paiho, S., Mäki, E., Wessberg, N., Paavola, M., Tuominen, P., Antikainen, M., Heikkilä, J., Rozado, C., & Jung, N. (2020). Towards circular cities—Conceptualizing core aspects. Sustainable Cities and Society, 59, 102143. https://doi.org/10.1016/j.scs.2020.102143

Pozzetto, E., & Leoni, L. (2025). Communication of circular economy in the public sector in Bulgaria: A study of space, agenda, language, and tools. Circular Economy and Sustainability, 5(6), 5727–5748.

Ranta, V., Aarikka-Stenroos, L., & Mäkinen, S. J. (2018). Creating value in the circular economy: A structured multiple-case analysis of business models. Journal of Cleaner Production, 201, 988–1000.

Rapley, T. (2014). Sampling strategies in qualitative research. In The SAGE handbook of qualitative data analysis (Vol. 4, pp. 49–63). SAGE.

Serrat, O. (2017). Knowledge solutions: Tools, methods, and approaches to drive organizational performance (p. 1140). Springer Nature.

Shah, P., & Yang, J. Z. (2025). Decoding derogation: The impact of environmental values and political ideology on the effect of persuasive message about recycle and reuse behaviors. Risk Analysis, 45, 1428–1442.

Termeer, C. J., Dewulf, A., & Biesbroek, G. R. (2017). Transformational change: governance interventions for climate change adaptation from a continuous change perspective. Journal of Environmental Planning and Management, 60(4), 558-576.

Turin Urban Innovators Guide. (2025). https://torinocitylab.it/storage/2025/10/TAG_Turin-Urban-Innovators-Guide_2025.pdf

Yin, R. K. (2009). Case study research: Design and methods (Vol. 5). Sage.

Yin, R. K. (2018). Case study research and applications. Sage.

Wynne, B. (1992). Misunderstood misunderstanding: social identities and public uptake of science. Public understanding of science, 1(3), 281.

Sustainability and Project Management: Discovering Skills Relations to Support Knowledge-Intensive Organizations

Irene Spada, Antonella Martini, Gualtiero Fantoni
University of Pisa, Pisa, Italy

Abstract

The integration of sustainability in project management is becoming essential to navigate the transformation of industries and societies overcoming the technical-environmental trade-off in Knowledge Intensive Organizations. Although sustainability is a key driver for long-term orientation and competitive advantage, identifying and operationalizing sustainability-related capabilities in project contexts remains complex. One challenge lies in the identification of skills and practices to understand how sustainability related capabilities can be operationalized in project contexts. This contribution propose a data-driven approach for skills paths design to map relations among skills with a focus on project management and sustainability requirements. This study considers field-specific requirements with over 5,000 PM-related job vacancies and Project Management Body of Knowledge (PM-BOK®), and adopts the European Classification of Skills, Competences and Occupation (ESCO). The methodology develops the so-called *Prerequisite Discovery* approach, that integrates text mining, graph theory, and machine learning, leveraging BERT-based word embeddings and link prediction algorithms, to extract and model skill interdependencies. Results identify 165 PM skills, including five green PM skills, namely environmental engineering, lean manufacturing, energy efficiency, ecosystems, and detecting bottlenecks. By mapping more than 14,000 prerequisite relations and 456 cross-domain links the study reveals the current PM-sustainability interplay. Project management competences can act as enablers for green skills, and conversely, how green skills function as foundational prerequisites for sustainability-oriented project practices. While the current PM-sustainability interplay remains predominantly technical, emerging connections with digital, behavioural, decision-making, and creative competences suggest the potential for a more holistic integration. The findings therefore indicate that PM green skills can serve as strategic enablers in knowledge creation, sharing, and integration. A framework is proposed to operationalize sustainability by grafting the knowledge cycle onto the green engine of the identified PM green skills. These information can feed adaptive educational and training paths to transform sustainability from a specialized concept into a cross-cutting strategic capability.

Keywords: skill, job vacancies, textbook, labour market intelligence

Paper type: Academic Research Paper

1 Introduction

Technological advancements in last decades are impacting every aspect of our lives (Wang et al., 2020) and are strongly intertwined with the increasing commitment towards sustainability (Saunila et al., 2019). Industry 4.0 and digitalization relate sustainability in multiple ways: improving performance and resources consumption optimization, but also with waste production and resource demand (Beltrami et al., 2021). In such context, managerial practices are gaining the momentum, as sustainable project management can be paramount to overcome the technical-environmental trade-off (Silvius & Schipper, 2014; Orieno et al., 2024). Companies are asked to reconfigure managerial competences, integrating technological skills and an holistic, systemic, and ethical approach, but guidance on how such competences relate to one another and how they can be effectively developed is still limited (Gomes et al., 2025). Training providers are proposing updated learning activities, although these efforts sometimes lag behind due to the different pace of labour-market evolution and educational programme renewal (Mukul & Buyukozkan, 2023; Tasdemir & Gazo, 2020). Data-driven methods can offer a powerful lever to timely inform educational and training design, enabling effective and efficient responses to technological and societal transformations (Ahmad et al., 2023; Nabizadeh et al., 2020).

The challenge thus is not only to integrate sustainability-oriented competences into project management programmes, but to understand which competences should be incorporated and how they interrelate within learning and career pathways. The reciprocal enabling relations between project management skills and sustainability-related capabilities, particularly in knowledge intensive contexts, has become a central managerial imperative for understanding how sustainability can be operationalized in project contexts. These considerations frame our Research Question as follows: *Which project management competences enable the development of sustainability, and which sustainability-related skills can advance project management capabilities relevant to sustainable project delivery?*

To address this two-side question, the study adopts a data-driven approach to map relations among technical, managerial, and sustainability-related competences in the field of Project Management. The relations

enable the mutual reinforcement of project management and sustainability capabilities. We develop *Prerequisite Discovery*, an approach combining text mining, graph theory, and machine learning to identify and model prerequisite skills relations. The approach adopts labour market data (job vacancies) and domain knowledge (Project Management Body of Knowledge PMBOK® 7th Edition and Green Skills from European Classification of skills and occupations ESCO). *Prerequisites Discovery* is deployed on 3 tasks: (1) *Competence Identification* extracts competences from job vacancies with their frequency (Kobayashi et al., 2018); (2) *Prerequisite Relation Identification* delineates skills prerequisites from domain related books based on rules of mentioning in the books (Xiao et al., 2022; Lu et al., 2019); (3) *Prerequisite Relation Prediction* defines relations among demanded exploiting word embeddings in link prediction algorithm (Devlin et al., 2018; Kumar et al., 2020). This enables the construction of a map that highlights competence relations supporting sustainability-oriented project management.

This study aims to advance design for sustainable project management training opportunities by uncovering the structure of relations among technical, managerial, and sustainability-related skills using a data-driven approach. The paper explores how project management competences can act as enablers for green skills, and conversely, how green skills function as foundational prerequisites for sustainability-oriented project practices. The map of skills describe the development of sustainable project management capabilities, presenting the configuration of technical and managerial competences required for project management roles together with the sustainable related skills. The structured representation of skills relations provides a robust foundational knowledge base that could be leveraged in large-language-model architectures to generate informed inputs for adaptive and sustainability-oriented training programmes. The findings reveals how green skills are embedded within PM competence structures, showing both the prerequisites needed to develop green skills and the PM competences that green skills enable.

2 Background Literature

Sustainability-oriented project management is being addressed in literature in a triple-bottom line perspective: economic, social, and environmental (Martens & Carvalho, 2017). The convergence of green-related skill sets with managerial practices is leveraged for addressing the tension between technical performance and environmental responsibility, with key capabilities including technical expertise, behavioural dispositions, and value-driven orientations (Beducci et al., 2024; Consoli et al., 2016). An holistic, systemic, and ethically grounded approach is gradually permeating project management practices (Silvius & Schipper, 2014; Orieno et al., 2024; Gomes et al., 2025). A growing body of research is re-examining professional roles through a greener lens. Lupi et al. (2022) conceptualise an updated engineer archetype by quantifying technological competences and their contribution to Sustainable Development Goals achievement. The professional profile required for the green transition extends well beyond technical and job-specific expertise, including, business management skills, soft skills, environmental awareness, and willingness-to-engage (Nikolajenko-Skarbalė et al., 2021). Translating these professional requirements into educational pathways requires targeted actions (Suciu & Nsulea, 2019) and curricula update (Spada et al., 2022). Yet these efforts often struggle to keep pace with labour market demands (Mukul & Buyukozkan, 2023). Data-driven approaches can provide timely evidences, enabling agile and effective responses (Ahmad et al., 2023; Nabizadeh et al., 2020).

Relations among skills in education are expressed as prerequisites, i.e., the prior knowledge needed to understand a course, and learning outcomes, which specify what students are expected to achieve and to what level (Resnick et al., 1973; Biggs, 2012). Their links develop learning paths, sequences of learning objects or courses (Nabizadeh et al., 2020). Such paths support study-plan design (Auvinen et al., 2014), career planning (Zhang et al., 2019; Sodanil et al., 2019), and curriculum tailoring for specific job roles (Demchenko et al., 2021). To model these relations, scholars increasingly rely on data-driven and AI-based approaches (Ahmad et al., 2023), such as binary classification (Liang et al.,2018; Xiao et al. 2022), multidimensional scaling (Ang et al., 2014), graphs and network analysis (Wang et al., 2016; Auvinen et al., 2024). Sequential learning materials, as textbooks and MOOCs, are exploited (Lu et al., 2019; Wang et al., 2016). Learning content can be connected as career trajectories, providing insights for career paths and detecting skill gaps for specific jobs (Sodanil et al., 2019; Demchenko et al., 2021).

3 Data and Methods

The method is deployed on 3 tasks (Figure 1). It starts with in-demand skills Identification from job vacancies, then prerequisite relations are mapped by analysing domain related books, and finally these information are used to predicted prerequisite links. The final map focus on the intersection of sustainability and project management.

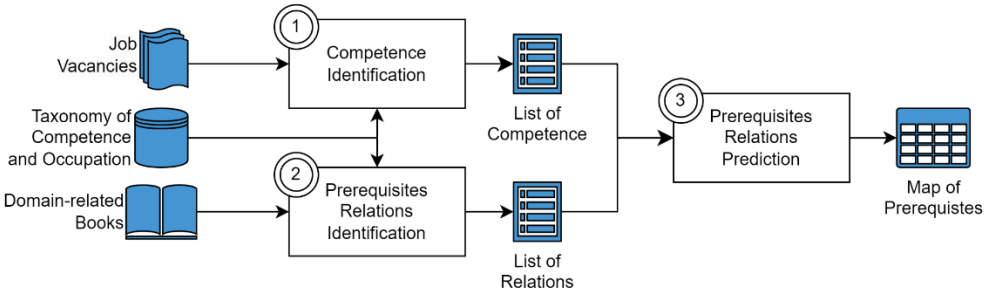


Figure 1 - Flowchart of the proposed method.

The reference framework for labour market terminology is the European Classification of Skills Competences and Occupations (ESCO). It is a multilingual classification of skills, jobs and qualifications, adopted in studies on labour market and education (Chiarello et al., 2021; Spada et al., 2022). It is based on the International Standard Classification of Occupations (ISCO-08) and continuously updated. Each skill provides a description, a tag for the type (e.g., skill/knowledge), and is related to one or more professional profiles. The same applies for job profiles. The English version v1.1 was used, including 13.890 skills and knowledge and 3.008 job profiles. ESCO includes also a green label for 571 items referring "knowledge, abilities, values and attitudes needed to live in, develop and support a society which reduces the impact of human activity on the environment" (Cedefop, 2012).

Job vacancies are the data for labour market demand, as widely used to analyse the competences and professionalism required and examine trends (Kobayashi et al., 2018).

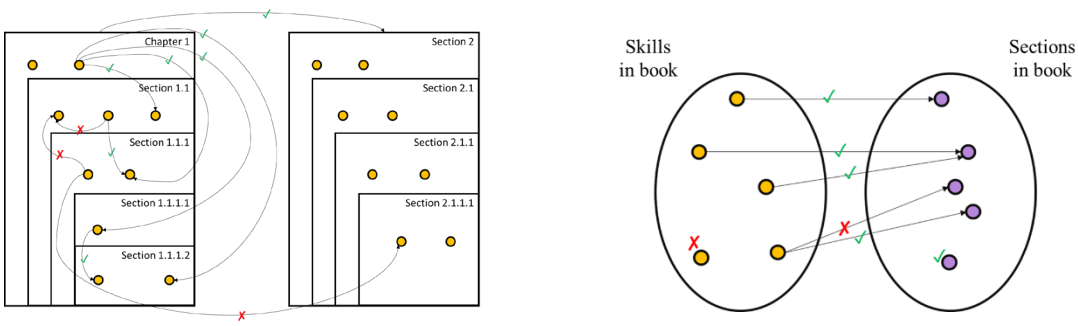
Books are used to describe domain requirements and relations among the learning concepts considering their table of content (Lu et al., 2019; Wang et al., 2016). We used the 7th edition of the PM-BOK® ("*A guide to the project management body of knowledge (PM-BOK® guide) and the standard for project management*" published by Project Management Institute in 2021) as it is widely credited by professional and institutional PM communities (Timinger et al., 2021).

3.1 Competence identification

Competences from ESCO are searched in job vacancies using a string-matching algorithm for exact matches. Then, the occurrence analysis in job vacancies, i.e., counting the number of ads mentioning a given competence, was used to identify the most relevant competences for project management (Kobayashi et al., 2018).

3.2 Prerequisites Relations Identification

The domain-related competencies are searched in books using string-matching algorithms with ESCO, as for the job vacancies. The structure of the book is processed following its table of contents. Each of the identified competences is assigned to only one section considering the order of mention in the book. This hypothesis is based on literature on hierarchical task relations (Nabizadeh et al., 2020; Resnick et al., 1973). From this hypothesis we derive the assumptions about competence-section belonging and ranking, presented in Figure 2. For the *assumption of belonging*, each competence can be mentioned in more than one section, but one and only one section elaborates on it; each section can mention one, several, or no one competences. For the *assumption of ranking*, the relation can exist only inside a single chapter in subsequent sections or in different chapters considered as a whole.



(a) *(b)*

Figure 2 - Association and positioning rules in assumptions of belonging (a) and ranking (b). Green marks (✓) are for the acceptable cases, red marks (✗) otherwise.

3.3 Prerequisites Relations Prediction

We combine the obtained information in the machine learning process in Figure 3. Competencies identified in book are used to train a neural network in predicting prerequisite links. Competences identified in ads but not in book have been processed with the trained model to identify new links. The predictive features include: the ESCO tag for skill and knowledge concepts (ESCO Skill-Type, https://esco.ec.europa.eu/en/escopedia/skills-pillar), the ESCO metric for degree of competence reusability, i.e., transversal, cross sectoral, sector specific, and occupation specific (ESCO Reuse Level, http://esco.ec.europa.eu/en/escopedia/skill-reusability-level), and the vectorized representation of competence labels obtained using BERT model (Devlin et al., 2018). We used the full ESCO classification to map the overall landscape and its relationships, while the ESCO Green Collection was used to examine the results on how sustainability relates with project management.

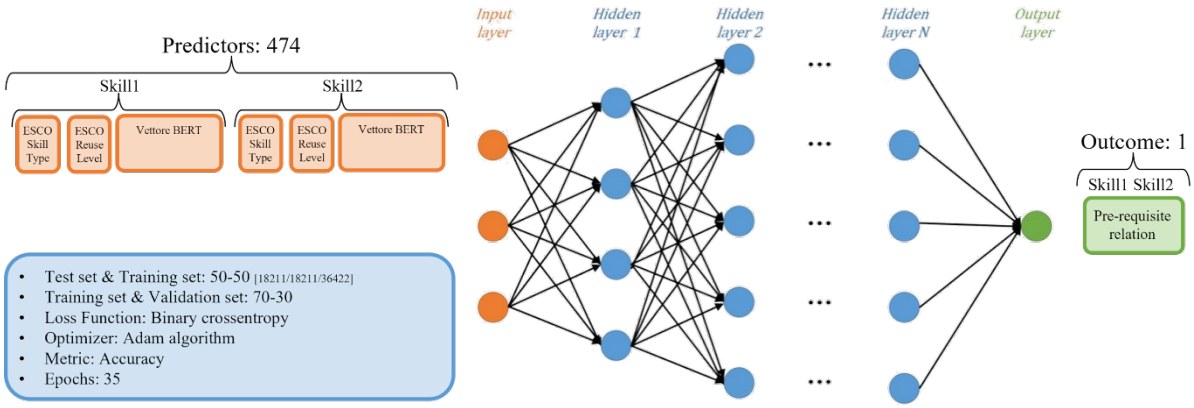


Figure 3 - Machine model used to predict the links among the competences.

4 Results and Discussion

We analysed 5,885 LinkedIn job vacancies, published between August 2021 and February 2022. We identify: 165 relevant PM competences, five of which tagged as *green skill* in ESCO Green Collection; 5,535 prerequisite relationships among 116 competences within PM-BOK®; 8,488 new prerequisite relations with a probability threshold above 70%[1]. These relations were enriched with metadata from ESCO, including skill types, reuse levels, and relevance based on job market demand. The ESCO Green Collection allows us to identify 103 relations where green skills serve as prerequisites of PM skills and 353 relations where PM skills enable green capabilities.

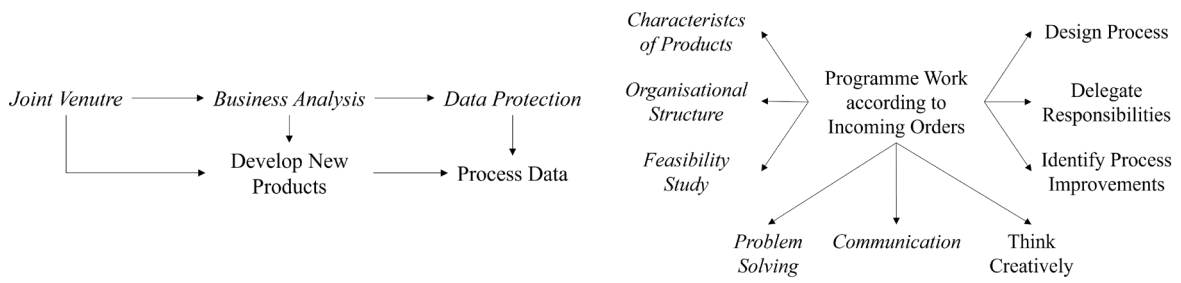
4.1 Overview from job vacancies and PM-BOK®

Figures 4 show the charts of the top 30 most requested competences resulted from job ads. The two rankings provide an overview on the core project management knowledge and skills, fundamental for effective project execution, team coordination, and project planning. Analytical and decision-making abilities characterise this professional field. But also digital and technical expertise emerges, more and more important in the contemporary digital business environment, as well as transversal abilities related to creativity and adaptability. Overall, these lists illustrate the multifaceted nature of project management and collectively shape the complexity of the project management role.

Figure 5 presents two examples of relations from the PM-BOK® field. The skills path in presents how theoretical foundations (e.g., *business analysis*, *data protection*) support new product development and data processing; while the example in Figure 7b shows how practical planning abilities can support effective workflow design, stakeholder alignment, and adaptive problem-solving.

[1] *The machine learning algorithm demonstrated strong performance metrics (Precision 0.931, Recall 0.904, F1 score 0.918, Accuracy 0.916), predicting 22,900 potential skill links, resulting in almost 8,500 relations based on the set threshold. We used: feedforward artificial neural network (ANN) via keras package in R, TensorFlow backend, Adam optimizer and binary cross-entropy loss function. Model performance was evaluated using typical metrics (Puccetti et al., 2023). Training and test sets were in 50% split proportion.*

Figure 4 - Top 20 ESCO knowledge and skills extracted from job vacancies in the PM field.



(a) *(b)*

Figure 5 - Skills path of 'Joint venture' (a) and 'Programming work according to incoming orders' (b) in PM-BOK®. Skills are in plain text, knowledge in italic.

4.2 Map of sustainability-oriented project management

The full *Map of Prerequisites* for the PM field is available in interactive version (*link_will_be_available*). Here the intersection between PM and sustainability are reported.

Among the 165 ESCO competences identified in the PM field, only five are tagged as green. This suggests that the integration between PM and sustainability is still in the rough: there is a gap between the ambitions and the reality of the labour market. Those five competences are strongly connected, as depicted in Figure 6: environmental engineering is the theoretical base enabling two direction, i.e., energy efficiency and lean manufacturing. Thus, sustainability is intended in its engineering lens of process optimization to improve environmental impact and waste reduction. An attention to a more systemic approach can be suggested by ecosystems, however, this is still a too weak signal to be confirmed. So green skills are few, fragmented and technical, and sustainability is seen as an efficiency goal.

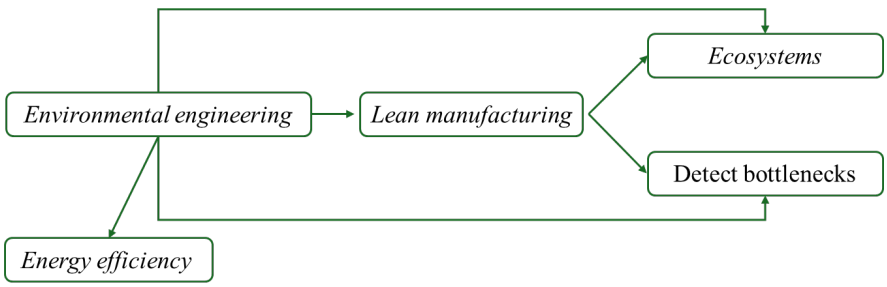


Figure 6 - Skills relations among green PM skills. Skills are in plain text, knowledge in italic.

These green PM skills can be enabled by PM skills or be prerequisites. In the first case, we can measure how expensive is acquiring green skill, so their cognitive dependence. In the second case, we can measure how much green PM skills contribute to PM transformation. Table 1 shows these values. Green PM skills require many prerequisites, possibly because they are considered advanced and specialized. These values confirm sustainability is not yet a transformative driver of PM, because only 2 skills are prerequisites of other PM skills, and one of these (lean manufacturing) is more lean than green. In this overview, environmental engineering appear as the only "cognitive hub", connecting a substantial group of skills.

Table 1 – Number of PM skills required as prerequisites for green skills and enabled by green skills.

	Green PM Skills				
	Environmental engineering	**Lean manufacturing**	**Detect bottlenecks**	**Ecosystems**	**Energy efficiency**
N. of PM prerequisite	45	59	92	92	75
N. of PM skill enabled	77	20	0	0	0

Figures 7 and 8 represent skills as nodes and their relations as flows, coloured to distinguish PM skills and green PM skills. The structure is organised into different vertical stages with prerequisites and enabled skills from left to right. Figure 7 presents the most in demand PM skills, Figure 8 reports PM-related soft skills; both map their links to the five green PM skills

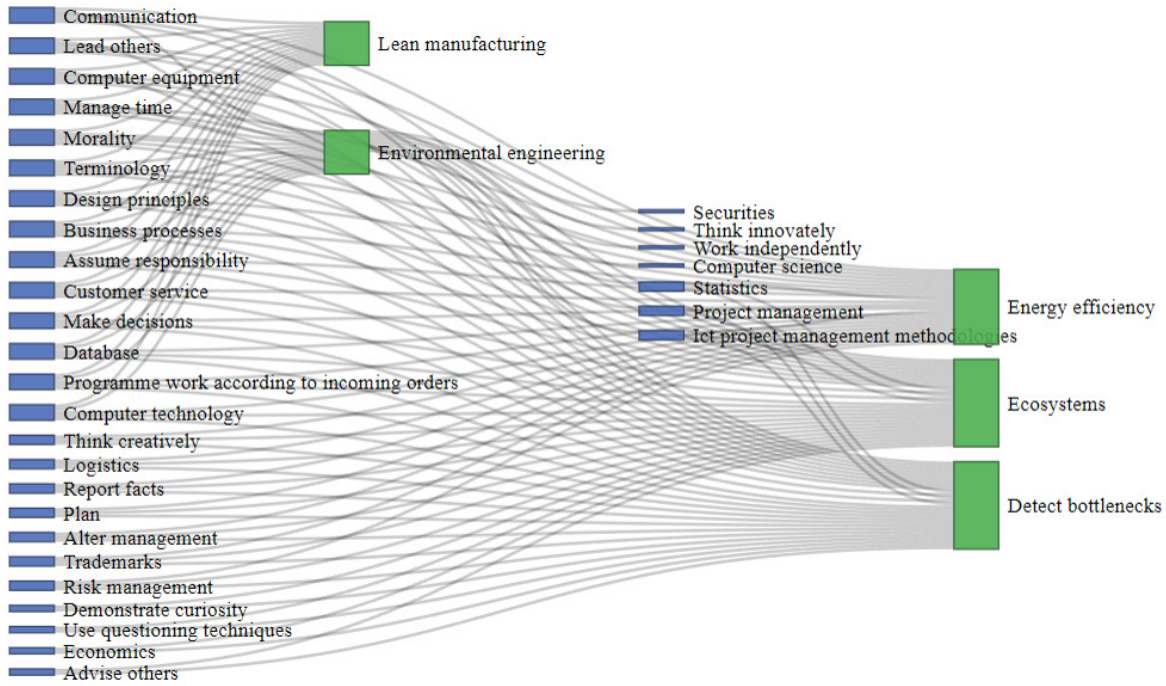


Figure 7 - Diagram of relations among PM skills, in blue, and green PM skills, in green.

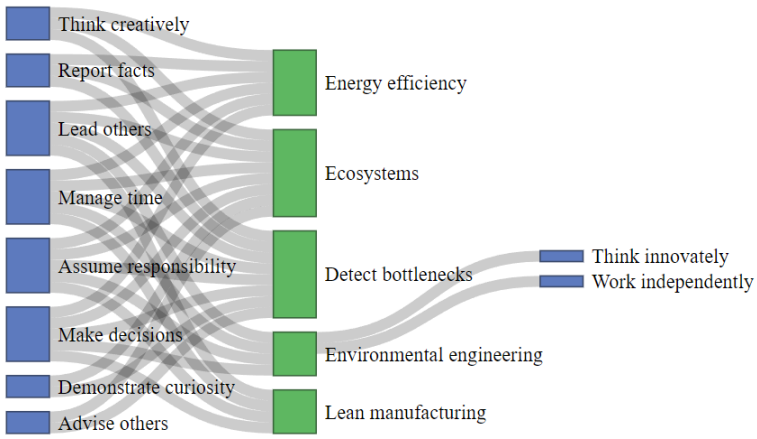


Figure 8 - Diagram of relations among soft skills related to PM, in blue, and green PM skills, in green.

In Figure 7, PM green skills emerge as convergence nodes and are highly dependent on traditional PM skills: the flow from left to centre is denser than the flow from centre to right. A technical, engineering, efficiency-oriented sustainability emerges, given the presence of skills such as *design principles*, *business process*, *programme work according to incoming orders*, *logistics*. There are digital skills (*computer technologies*, *database*, *computer science*), decision-making and behavioural abilities (*lead others*, *assume responsibility*, *advise others*), as well as *think innovatively* and *creatively*; suggesting data-driven reasoning can support green practices, particularly for optimisation and system integration. Sustainability-related approach is highly dependent on technical expertise but also on interpersonal competence and adaptive mindset, combining in a hybrid profile engineering, digital literacy, decision-making abilities, and creative thinking.

Figure 8 deepen the analysis only on soft skills. In this diagram sustainability is enabled by behavioural and cognitive abilities, emphasizing the role of transversal skills act as enablers for a greener approach, supporting project managers to engage with sustainability challenges and not just comply with standard and optimization.

4.3 Interplay between PM and sustainability skills in KIOs

Knowledge is the primary resource and the main source of competitive advantage for Knowledge Intensive Organizations (KIOs) (Makani and Marche, 2010). KIOs have flexible organizational structure, with unique and not predefined processes, and highly specialize human capital, with high discretionary power; usually coordination and exchanges among experts does not follow pre-established mechanisms (Makani and Marche, 2010). These elements often make traditional PM paradigms ineffective, especially with standardized work plans (Nieto-Rodriguez A., 2021). Sustainability introduces additional peculiarities that can radically transform the way these organizations operate (Bertello et al 2022). The simultaneous integration of multidimensional objectives on long term brings out the need to define programmatic objectives and plans. So PM capabilities became paramount: project managers should facilitate knowledge creation, integration and sharing, ensuring the required flexibility while operationalizing sustainable capabilities in knowledge-driven environments (Lundin et al., 2015).

The present study depicts how sustainability-related skills are remains predominantly technical and efficiency-oriented within the broader architecture of PM skills. The support of transversal and digital competences appear still not embedded as a strategic capability. For KIOs, where sustainable project delivery requires not only technical proficiency but also systems thinking, these results further emphasize the need of reconfiguring PM skills for integrating sustainability not as a specialized add-on but as a cross-cutting capability. Such integration can shape knowledge practices to strengthen the whole project lifecycle for sustainable knowledge transformation cycle.

Knowledge creation promoting sustainable oriented experimentation and practices (e.g., Medina and Medina, 2015; Sjödin, 2019). *Environmental engineering* and *energy efficiency* enable continuous exploration of technical and regulatory solutions, feeding experimental learning cycles. *Lean manufacturing* and *detect bottlenecks* promote rapid experimentation, iterative problem-solving, and continuous improvement, all of which are central to knowledge creation processes. Those abilities also push cross-functional teams to identify inefficiencies and co-design innovative solutions, and integrate operational, analytical, and digital approaches, facilitating alignment between efficiency, quality, and sustainability goals.

Knowledge sharing leveraging cross-functional collaboration and communities of practice (e.g., Yeh, 2008; Barbosa and Carvalho, 2025). *Lean manufacturing* favors the dissemination of good practices and lessons learned across projects, functions and departments. *Ecosystems* introduces a systems perspective that encourages interdependencies.

Knowledge integration adopting frameworks that connect processes and competences (e.g., Atkočiūnienė et al., 2022; Oyeyipo et al., 2026). *Environmental engineering* and *ecosystems* involve technical experts, environmental stakeholders and project managers, creating spaces for structured exchange, and help to connect PM skills with sustainability and impact assessment logics.

From these considerations, we derive the visual framework in Figure 9, where green PM skills are positioned as enablers of organizational change, supporting knowledge creation, sharing, and integration in KIOs. When green PM skills are considered as strategic enablers, rather than mere add-ons, they provide the necessary structure through which sustainability-related capabilities can be operationalized within project contexts as a green engine. This process of operationalization ensures that the knowledge cycle in KIOs can be effectively grafted onto the organization's practices. The green PM skills can thus be drivers for continuous experimentation (creation), cross-functional dissemination of best practices (sharing), and alignment of diverse technical expertise (integration), leveraging sustainability as a cross-cutting capability in project lifecycle.

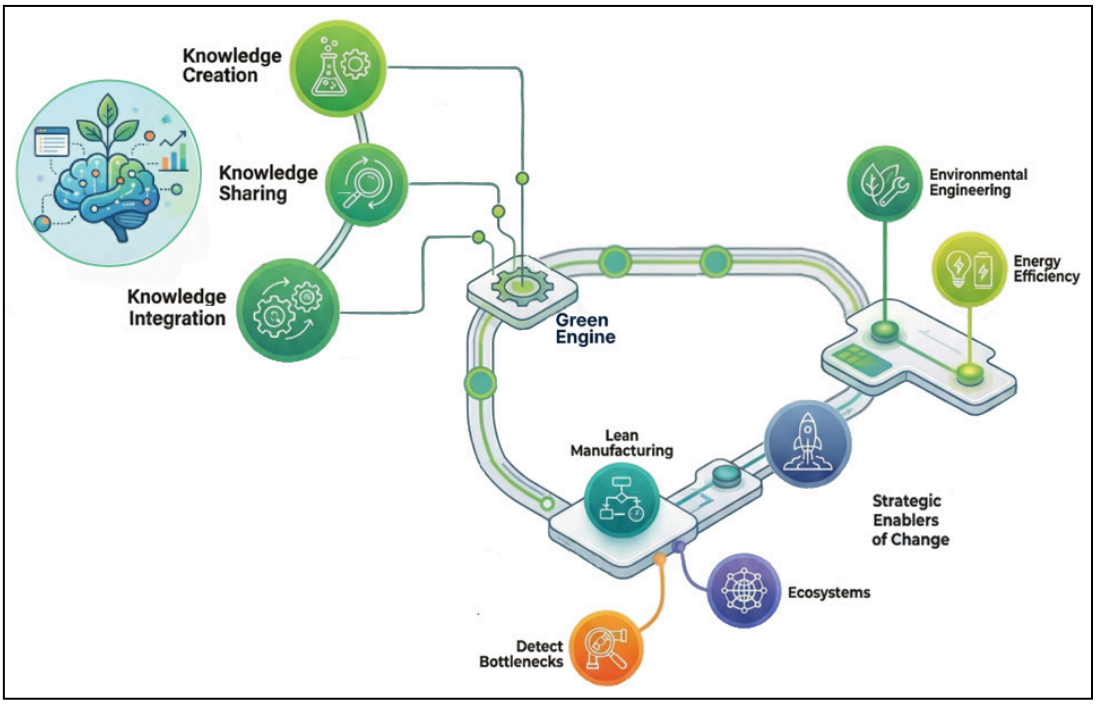


Figure 9 - Framework of PM green skills as enablers for sustainable knowledge transformation cycle.

6 Conclusions

The study extracts in-demand skills from over 5,000 PM-related job vacancies, maps prerequisite relations from PM-BOK®, and predicts additional links, identifying 165 PM skills including five green skills, over 14,000 prerequisite relations, and 456 cross-domain links showing how PM and sustainability relate. Results emphasize a technical and lean-centric approach in PM-sustainability interplay. The relations with digital skills, decision-making and behavioural abilities, as well as creativity and soft skills suggest that a more holistic approach can be reached.

The proposed approach includes market-driven data, prioritizing the logical sequence of skills development rather and job profile/roles needs. This ensures that learners and professionals can develop expertise through a structured progression. These results offer actionable guidance for designing training programmes that support the evolution of skills for sustainable project management in knowledge-intensive organizations, fostering knowledge integration, cross-functional collaboration, and sustainability-oriented decision-making. We leverage on the established techniques in the field of educational design (graph-based machine learning) for identifying skills relations. This enable potential for application across various educational contexts, including universities, schools, and training providers for upskilling and reskilling initiatives, support programmes update and lifelong

learning. Educational institutions and training providers can improve their responsiveness to labour market trends, enhancing employability and relevance.

We must acknowledge that when skills are anchored within a specific domain (like PM in the analysed case) the skills requirements may be constrained by the field knowledge and practices. Indeed, the amount of green skills detected is strongly dependent on the set of data adopted (job vacancies and PM-BOK®). This raises awareness on the importance of keeping as much as open the design process, allowing the integration of skills set from various disciplines.

The work can be improved, paving the way for future studies. Concerning the data, here just one textbook is used, but the several domain-specific books can feed the process, contributing to a more fine-grained result and probably positively affecting the performance. It is possible to use different techniques for identifying skills. Such may include other Named Entity Recognition approaches and also the novel Generative AI techniques. Then, further tests can be executed on which parameters of the prediction algorithm contribute at most for higher performance of the model. Future studies can explore how PM green skills can be integrated in organizational culture and knowledge-sharing practices. Sustainability norms and PM practices can vary across industries and countries, so comparative researches can identify sector-specific patterns, maturity levels and gaps, cultural and institutional factors influencing skills integration, offering more targeted insights for training and policy. Empirical studies could assess further effects of integrating sustainability and PM on company performance, stakeholder engagement, innovation, strengthening sustainability strategies.

7 Acknowledgements

Authors would like to thank Beam Me Up (a spin-off of the University of Pisa), for the possibility to test and implement the proposed method in real use cases.

8 Ethics Declaration

The study did not involve people or animals, thus ethical clearance was not required.

9 AI Declaration

The authors declare that NotebookLM was used for generating icons and visual elements included in the framework visualization (Figure 9). All components were subsequently assembled, reviewed, and refined by the authors. The authors take full responsibility for the accuracy and originality of the work.

References

Ahmad, K., Iqbal, W., El-Hassan, A., Qadir, J., Benhaddou, D., Ayyash, M., & Al-Fuqaha, A. (2023). Data-driven artificial intelligence in education: A comprehensive review. *IEEE Transactions on Learning Technologies*.

Ang, L., D'Alessandro, S., & Winzar, H. (2014). A visual-based approach to the mapping of generic skills: its application to a Marketing degree. *Higher Education Research & Development*, *33*(2), 181-197.

Auvinen, T., Paavola, J., & Hartikainen, J. (2014, November). STOPS: a graph-based study planning and curriculum development tool. In *Proceedings of the 14th Koli Calling International Conference on Computing Education Research* (pp. 25-34).

Atkočiūnienė, Z. O., Gribovskis, J., & Raudeliūnienė, J. (2022). Influence of knowledge management on business processes: value-added and sustainability perspectives. *Sustainability*, *15*(1), 68.

Barbosa, M. T., & Carvalho, M. M. (2025). Enhancing sensing capabilities in project-based organizations: the role of knowledge transfer between commercial and operational teams. *International Journal of Project Management*, *43*(5), 102724.

Beltrami, M., Orzes, G., Sarkis, J., & Sartor, M. (2021). Industry 4.0 and sustainability: Towards conceptualization and theory. *Journal of Cleaner Production*, *312*, 127733.

Biggs, J. (2012). *Enhancing learning through constructive alignment*. Enhancing the Quality of Learning: Dispositions, Instruction, and Learning Processes.

Cedefop. (2012). Green skills and environmental awareness in vocational education and training: Synthesis report.

Chiarello, F., Fantoni, G., Hogarth, T., Giordano, V., Baltina, L., & Spada, I. (2021). Towards ESCO 4.0–Is the European classification of skills in line with Industry 4.0? A text mining approach. *Technological Forecasting and Social Change*, *173*, 121177.

Demchenko, Y., Maijer, M., & Comminiello, L. (2021, February). Data Scientist Professional Revisited: Competences Definition and Assessment, Curriculum and Education Path Design. In *2021 4th International Conference on Big Data and Education* (pp. 52-62).

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

Jose Lopes Gomes, R., Ferreira da Silva, L., Rezende da Costa, P., & Goncalves de Oliveira, P. S. (2025). Digital technologies and knowledge management in project context: a systematic literature review. *Knowledge Management Research & Practice*, *23*(6), 749-764.

Kobayashi, V. B., Mol, S. T., Berkers, H. A., Kismihók, G., & Den Hartog, D. N. (2018). Text mining in organizational research. *Organizational research methods*, *21*(3), 733-765

Kumar, A., Singh, S. S., Singh, K., & Biswas, B. (2020). Link prediction techniques, applications, and performance: A survey. *Physica A: Statistical Mechanics and its Applications*, *553*, 124289.

Liang, C., Ye, J., Wang, S., Pursel, B., & Giles, C. L. (2018, April). Investigating active learning for concept prerequisite learning. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 32, No. 1).

Lu, W., Ma, P., Yu, J., Zhou, Y., & Wei, B. (2019). Metro maps for efficient knowledge learning by summarizing massive electronic textbooks. International Journal on Document Analysis and Recognition (IJDAR), 22, 99-111.

Martens, M. L., & Carvalho, M. M. (2017). Key factors of sustainability in project management context: A survey exploring the project managers' perspective. *International journal of project management*, *35*(6), 1084-1102.

Medina, R., & Medina, A. (2015). The competence loop: Competence management in knowledge-intensive, project-intensive organizations. *International Journal of Managing Projects in Business*, *8*(2), 279-299.

Mukul, E., & Büyüközkan, G. (2023). Digital transformation in education: A systematic review of education 4.0. *Technological Forecasting and Social Change*, *194*, 122664

Nabizadeh, A. H., Leal, J. P., Rafsanjani, H. N., & Shah, R. R. (2020). Learning path personalization and recommendation methods: A survey of the state-of-the-art. *Expert Systems with Applications*, 159, 113596.

Orieno, O. H., Ndubuisi, N. L., Eyo-Udo, N. L., Ilojianya, V. I., & Biu, P. W. (2024). Sustainability in project management: A comprehensive review. *World Journal of Advanced Research and Reviews*, *21*(1), 656-677.

Oyeyipo, O., Aiyamenkhue, E., Adeyi, A. S., Adewolu, A., & Fadeyi, A. A. (2026). Green strategies for sustainable project management practice: a pathway to sustainable cities and communities. *International Journal of Construction Management*, 1-17.

Puccetti, G., Giordano, V., Spada, I., Chiarello, F., & Fantoni, G. (2023). Technology identification from patent texts: A novel named entity recognition method. *Technological Forecasting and Social Change*, *186*, 122160.

Resnick, L. B., Wang, M. C., & Kaplan, J. (1973). Task analysis in curriculum design: A hierarchically sequenced introductory mathematics curriculum 1. *Journal of Applied Behavior Analysis*, 6(4), 679-709.

Saunila, M., Ukko, J., & Rantala, T. (2018). Sustainability as a driver of green innovation investment and exploitation. *Journal of cleaner production*, *179*, 631-641.

Silvius, A. J., & Schipper, R. P. (2014). Sustainability in project management: A literature review and impact analysis. *Social business*, *4*(1), 63-96.

Sodanil, M., Chotirat, S., Poomhiran, L., & Viriyapant, K. (2019, June). Guideline for Academic Support of Student Career Path Using Mining Algorithm. In *Proceedings of the 2019 3rd International Conference on Natural Language Processing and Information Retrieval* (pp. 133-137).

Sjödin, D. (2019). Knowledge processing and ecosystem co-creation for process innovation: Managing joint knowledge processing in process innovation projects. *International Entrepreneurship and Management Journal*, *15*(1), 135-162.

Spada, I., Chiarello, F., Barandoni, S., Ruggi, G., Martini, A., & Fantoni, G. (2022). Are universities ready to deliver digital skills and competences? A text mining-based case study of marketing courses in Italy. *Technological Forecasting and Social Change*, *182*, 121869.

Tasdemir, C., & Gazo, R. (2020). Integrating sustainability into higher education curriculum through a transdisciplinary perspective. *Journal of Cleaner Production*, *265*, 121759.

Timinger, H., Schmidtner, M., & Reiche, F. (2021). Comparison of different Standards and Norms in Project Management. *Research Notes on Data and Process Science*, (2), 1-7.

Wang, S., Ororbia, A., Wu, Z., Williams, K., Liang, C., Pursel, B., & Giles, C. L. (2016, October). Using prerequisites to extract concept maps fromtextbooks. In *Proceedings of the 25th acm international on conference on information and knowledge management* (pp. 317-326).

Wang, L., Luo, G. L., Sari, A., & Shao, X. F. (2020). What nurtures fourth industrial revolution? An investigation of economic and social determinants of technological innovation in advanced economies. *Technological Forecasting and Social Change*, *161*, 120305.

Xiao, K., Fu, Y., Deng, Y., & Xia, L. (2022, May). Identifying prerequisite relations between concepts in wikipedia. In *2022 International Conference on Service Science (ICSS)* (pp. 271-276). IEEE.

Yeh, H. (2008). A knowledge value creation model for knowledge-intensive procurement projects. *Journal of Manufacturing Technology Management*, *19*(7), 871-892.

Zhang, D., Liu, J., Zhu, H., Liu, Y., Wang, L., Wang, P., & Xiong, H. (2019, November). Job2Vec: Job title benchmarking with collective multi-view representation learning. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management* (pp. 2763-2771).

Mentoring as a Sociodigital Ecosystem: Knowledge Sharing, Digital Communities and Student Support in an Engineering Department

Demetrio Alessandro Trunfio[1], Ferdinando Potenza[1], Anna Mezzapesa[1], Carl Harvey Miguel[2], Giovanna Rotella[1]
[1] LUM University, Casamassima, Italy
[2] University of Bari, Bari, Italy

Abstract

Universities increasingly face the challenge of supporting diverse engineering students through critical transitions, while leveraging mentoring and digital infrastructures as strategic levers for knowledge management and student success. Building on mentoring and communities-of-practice literatures, this paper conceptualises mentoring in an engineering department as a sociodigital ecosystem that integrates formal, informal and peer-based practices with digital communities and AI-enabled tools. Empirically, the study focuses on the *"continuous mentoring"* initiative developed at the Engineering Department of a medium-sized private university in Southern Italy, where a department-wide digital community, cohort and event-based groups, course-specific chats and project teams have progressively evolved into a multi-layered knowledge infrastructure.

The research adopts a qualitative single-case design, drawing on internal project reports, departmental documents, social media analytics, extended abstracts and digital traces from student communities. These materials were coded inductively and organised into an ecosystem representation capturing layers, roles and knowledge flows, then interpreted through the lens of mentoring ecosystems and socio-digital knowledge infrastructures. The findings show how a persistent core community, surrounded by cohort-specific and project-based groups, supports first-year students by lowering the threshold for asking questions, sharing tacit tips and accessing timely guidance from faculty mentors, peer tutors and community administrators. The same ecosystem also amplifies participation in extracurricular and research-oriented initiatives, such as the *European Researchers Night* and the *Ingenium* entrepreneurship team, enabling students to mobilise situated know-how on how to prepare, present and learn from comparison with peers from other institutions.

Theoretically, the paper extends mentoring research by framing mentoring not as a single programme, but as an ecosystem of interconnected practices embedded in socio-digital infrastructures, where human relations, community practices and AI-enabled assistants jointly sustain knowledge creation and sharing. Practically, it identifies three design principles for universities seeking to build mentoring-based knowledge infrastructures: anchoring mentoring in a persistent core community; cultivating flexible satellite communities around cohorts, courses and projects; and augmenting, rather than replacing, human mentoring with AI tools for routine queries. The contribution speaks to scholars and practitioners interested in mentoring, knowledge management and the design of learning ecosystems in engineering and beyond.

Keywords: Knowledge management, mentoring, sociodigital skills, Communities of Practice, third mission

Paper type: Practical Paper

1 Introduction

Research on mentoring highlights multiple models – informal, peer, reverse and formal – as complementary mechanisms for supporting knowledge transfer, development and retention in organisations (Karkoulian et al., 2008; Short, 2014; Yusof et al., 2012). However, most empirical studies focus on workplace settings and pay less attention to how mentoring can be orchestrated as a continuous ecosystem for engineering students, where the risk of early drop-out and disengagement remains significant (Bryant, 2006; Fowler et al., 2019). This paper addresses this gap by analysing a "*continuous mentoring*" initiative at LUM's Engineering Department, examining how structured and peer-based mentoring practices, supported by digital and AI-enabled tools, contribute to knowledge sharing, student success and community building.

In this context, the Engineering Department at LUM University has developed a mentoring continuum that combines institutional mentoring, peer mentoring and informal community practices to support students from their first contact with the programme to the more advanced stages of their academic journey. The initiative specifically targets critical transition phases – such as the first year, changes of study path and entry into representative or project roles – by offering structured guidance, proximity support from senior students and digitally mediated access to knowledge resources. This mentoring ecosystem is embedded in a broader strategy of student engagement and retention, positioning mentoring not as a one-off intervention, but as a permanent knowledge infrastructure that connects people, practices and digital tools across cohorts.

The contribution of this paper is twofold: first, it conceptualises the LUM Engineering initiative as a "mentoring ecosystem" that integrates formal, informal and peer-based mentoring with socio-digital infrastructures for knowledge sharing; second, it offers an empirically grounded account of how such an ecosystem can be designed and managed to reduce fragmentation, support transitions and enhance student success. Accordingly, the study is guided by the following research questions: (1) how do different forms of mentoring formal, informal, peer contribute to knowledge creation and sharing in an engineering education context? (2) in what ways do digital and AI-enabled tools augment or transform mentoring practices within this ecosystem? and (3) what design principles emerge for universities seeking to build mentoring-based knowledge infrastructures that are scalable and sustainable over time?

2 Theoretical Background

Mentoring has long been recognised as a relational mechanism that supports both individual development and organisational knowledge processes, especially where tacit, experience-based knowledge plays a central role (Karkoulian et al., 2008; Yusof et al., 2012). Prior studies distinguish between formal mentoring, structured and institutionally designed, and informal mentoring, which emerges spontaneously from everyday work relationships; evidence suggests that informal mentoring may have particularly strong effects on knowledge sharing in contexts where trust and social proximity are critical (Karkoulian et al., 2008; Bjørnson and Dingsøyr, 2005). More recent work extends this view by emphasising peer mentoring among colleagues at similar hierarchical levels, showing that such arrangements can significantly predict organisational knowledge creation and sharing in knowledge- intensive settings such as software firms (Bryant, 2006).

Beyond informal and peer arrangements, the literature also highlights the role of structured mentoring programmes and reverse mentoring in shaping knowledge flows and innovation. Formal mentoring programmes, typically coordinated by HR and targeted at specific groups such as newcomers or high-potentials, have been associated with improved workplace networks, organisational outcomes and talent development, although their effectiveness depends heavily on the quality of the mentor–mentee relationship and programme governance (Srivastava, 2015; Fowler et al., 2019). At the same time, reverse mentoring initiatives, in which younger or more digitally skilled employees mentor senior colleagues, show how intergenerational exchange can accelerate the diffusion of digital capabilities, challenge traditional hierarchies and foster bidirectional learning in public and corporate settings (Short, 2014; Fowler et al., 2019). These contributions suggest that mentoring should be understood less as a single practice and more as a portfolio of interconnected mechanisms that, if designed as an ecosystem, can support knowledge sharing, innovation and inclusion across roles, cohorts and generations.

In parallel with mentoring research, work on communities of practice has shown how sustained, informal interaction around shared problems can generate powerful knowledge networks that complement formal structures (Wenger, 1998). In the LUM Engineering case, for instance, the growth of a large WhatsApp community around courses, exams and student initiatives has effectively turned everyday exchanges into a living CoP, where questions, solutions and tacit tips circulate continuously across cohorts. Digital group spaces, when combined with mentoring roles and light governance, can thus act as socio-technical infrastructures that make mentoring more scalable and visible, while preserving the relational richness of peer support.

Building on the mentoring and community-of-practice literatures, this paper adopts the notion of a "mentoring ecosystem" to capture the interdependence between multiple mentoring forms, community practices and socio-digital tools. Rather than treating mentoring programmes, peer support and digital communities as separate interventions, the ecosystem lens emphasises how they jointly constitute a knowledge infrastructure that spans roles (students, staff, administrators), temporal horizons (onboarding, progression, transition to work) and technological layers (messaging platforms, institutional systems, AI-enabled assistants). This perspective reson- ates with broader work on entrepreneurial and innovation ecosystems, but relocates the focus from firm–region dynamics to the micro-level of a university engineering department, where mentoring is mobilised as a strategic lever for talent development, retention and spin-off creation (Autio et al., 2018; Etzkowitz and Leydesdorff, 2000; Trunfio, 2024).

Table 1: Type of source and role in the study

Type of source	Example documents	Role in the study
Internal project reports and course outputs	"*Mentorship and Sharing Knowledge – The Continuous Mentorship*" (Design Thinking project); "*Incoming Mentoring*" slides	Reconstruction of mentoring practices, typology of mentoring models, identification of strategies and perceived outcomes.
Departmental documents	Engineering Department profile; mentoring project notes; social media performance report	Contextualisation of the case (programme characteristics, challenges, strategic objectives), description of digital channels.
Conceptual and empirical literature on mentoring and KM	Süveges and Kurucz (2024); Karkoulianet al. (2008); Bryant (2006); Srivastava (2015); Fowler et al. (2019); Short (2014); Yusof et al. (2012); Bjørnson and Dingsøyr (2005).	Theoretical framing of mentoring models and their relationship with knowledge management and knowledge sharing.
Literature on communities of practice and socio-digital KM	Romaldi (2002) (tacit–explicit via digital tools), AlAgha (2002) (Web 2.0, concept mapping and annotation), Hall (knowledge codification)	Framing of CoP, socio-digital infrastructures and codification/logics of knowledge transfer.
AI and KM materials	Trunfio(2024), Knowledge Management e l'AI Generativa – L.L.M. Base Solutions	Conceptualisation of AI- and RAG-based assistants as part of the mentor- ing ecosystem's knowledge infrastructure.

3 Case Context: The Engineering Department and its Mentoring Initiative

The empirical setting of this study is the Engineering Department of LUM University, an Italian private institution that has invested in positioning engineering education as a laboratory for innovation and student-centric learning. The department offers a management engineering programme with relatively small cohorts, which creates both opportunities for close interaction and challenges related to visibility, attraction and the risk of early drop-out in the first academic year. Over the last three years, the department has progressively developed a "*continuous mentoring*" initiative that integrates formal mentoring structures, peer-based support and digitally mediated communication channels, with the dual objective of improving onboarding and leveraging student com- munities as strategic assets.

3.1 Data collection

The study adopts a qualitative single-case design, combining multiple sources of evidence to reconstruct the mentoring ecosystem and its evolution over time. Data collection drew on: (1) internal documents and project reports on mentoring and student engagement, including the departmental mentoring project, design thinking outputs and social media analytics; (2) teaching and project materials produced in courses directly linked to the mentoring initiative; and (3) digital traces from the main communication and community channels used by engineering students, such as messaging groups and online platforms. These materials were complemented by informal conversations with key actors (programme coordinators, student mentors, community administrators), documented in analytic memos rather than formal interviews, according to the exploratory focus of the paper.

In addition to these documentary sources, we extracted basic activity metrics from the main digital communities used by engineering students, including the number of members per group, the share of active participants and the monthly volume of messages over a multi-year period. These metrics were not used for statistical modelling, but rather to calibrate the qualitative reconstruction of the mentoring ecosystem's layers and to distinguish between core, cohort-specific and project-based communities in terms of size and intensity of use. Finally, a preliminary semantic scan of message content was conducted using an external keyword analysis tool, with the aim of identifying recurrent terms and themes at community level. This semantic layer is not the focus of the present paper, but serves as a first step towards more systematic analyses of how concerns, emotions and topics evolve across the mentoring ecosystem over time. Figure 1 summarises the analytical flow adopted in this study. The paper combines documentary evidence, descriptive community metrics and a preliminary semantic scan of message content to reconstruct the mentoring ecosystem as a layered socio-digital infrastructure. This is not a full computational analysis, but a lightweight interpretive pipeline designed to contextualise the case, identify recurrent patterns and support the qualitative discussion of the ecosystem's functioning.

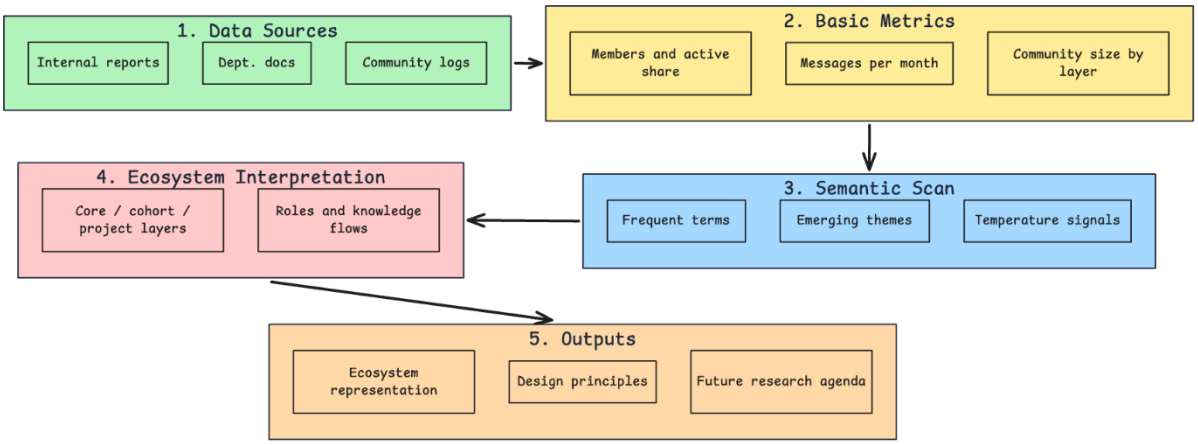


Figure 1: Analytical flow used to reconstruct the mentoring ecosystem through documentary sources, basic community metrics and preliminary semantic scanning.

3.2 Case context: LUM Engineering Department

The case is situated in the Department of Engineering of a medium-sized private university in Southern Italy, which offers undergraduate and graduate programmes in management and computer engineering. In recent years, the department has experienced a steady increase in enrolments and a progressive diversification of the student body, with a growing share of first-generation students and commuters. These dynamics have intensified typical challenges of engineering education, including high perceived difficulty of core courses, risk of early disengagement in the first year and the need to coordinate multiple actors involved in teaching, tutoring and student services.

To address these issues, the department launched a *continuous mentoring* initiative in 2023, designed as a multi-layered support system accompanying students from their first contact with the degree programme to graduation. The initiative combines formal elements – such as designated faculty mentors, structured tutoring activities and recognition of mentoring engagement within the study plan – with informal, student-led practices that emerged organically over time. A distinctive feature of the case is the deliberate use of digital communities and communication channels as the backbone of the mentoring system, turning scattered chats and groups into a more coherent sociodigital ecosystem.

At the core of this ecosystem lies a department-wide community that brings together students from different cohorts, peer tutors and faculty members in a persistent digital space. Around this core, additional communities have developed over time, including cohort-specific groups for first-year students, event-based communities such as those created for the annual European Researchers Night (ERN) initiative, and project teams such as the Ingenium team participating in entrepreneurship competitions. Course-specific chats, often initiated by students, represent a further layer focused on the day-to-day management of individual modules. Taken together, these communities form a dense network of interactions in which informational, emotional and knowledge-sharing exchanges coexist.

This configuration makes the department a particularly suitable case for examining mentoring as a sociodigital ecosystem. The coexistence of a stable core community, recurring event-based groups and temporary course or project teams offers a natural laboratory to compare different patterns of participation, content and value creation. It also allows the study to respond directly to the reviewers' request for a richer description of the context and of the specific characteristics that make the case relevant for the conference themes.

4 Layers and Actors of the Mentoring Ecosystem

The *continuous mentoring* initiative at the Engineering Department can be understood as a multi-layered ecosystem composed of partly overlapping communities, roles, and tools. At the centre lies a core department-wide community, hosted on a persistent messaging channel, where students from different cohorts, peer tutors, and faculty members interact around general announcements, cross-cutting questions, and community life. This core space functions as a shared "backbone" that provides visibility, quick access to help, and an always-on entry point into the mentoring ecosystem.

Surrounding this core, a second layer is formed by cohort-specific and event-based communities, including dedicated groups for first-year students, communities created for initiatives such as the annual ERN, and temporary teams such as the Ingenium group participating in entrepreneurship competitions.

Table 2: Overview of the mentoring ecosystem communities (March 2026)

Layer / function	Example communities	Typical size	Main role in the ecosystem
Core backbone	Main departmental noticeboard	177 members	Cross-cohort coordination, announcements, low-threshold entry point
Cohort groups	1st year bachelor engineering(L9); 2nd and 3rd year bachelor cohorts; 1st and 2nd year master's in management engineering	20 – 45 members	Navigation of the programme,first-year support, peer Q&A
Course communities	Quality management; Digital law & cybersecurity; Advanced robotics; Innovation & IP; Business eth- ics; Digital marketing; Risk management; Digital twin; Cloud data management	9 –18 members	Day-to-day management of modules, sharing of notes, tips and explanations among peers
Project and event groups	*ERN* 2025; Ingenium operations and events; Internship circle; Graduation sessions	4 –27 members	Preparation for events, project coordination, external visibility
Operational /technical	Erasmus engineering; iMensa technical core	3 –8 members	Logistical and technical support for mobility and lab-related activities

These communities focus on transition moments and collective projects, where mentoring takes the form of guidance on how to navigate the programme, coordinate activities, and translate opportunities into concrete experiences. A third layer is made up of course-specific chats and project groups, often initiated and moderated by stu- dents, which support the day-to-day management of individual modules, assignments and exams. Here, peer mentoring is particularly visible, as students share explanations, resources and tacit tips about how to approach specific topics, teachers or assessment formats. Across these layers, different actors play complementary roles – faculty mentors, peer tutors, community administrators and, increasingly, AI-enabled assistants that answer frequently asked questions – collectively sustaining the flow of informational, emotional and knowledge-rich exchanges within the ecosystem.

4.1 Knowledge flows in first-year support and extracurricular participation

Within the ecosystem, first-year students represent a critical target group, as they face the transition from secondary school to university, often with limited prior exposure to engineering programmes and with commuting or first-generation backgrounds. Knowledge flows towards this group through multiple, partially overlapping channels: core community announcements and FAQs, cohort-specific groups that address practical and emotional concerns, and course-specific chats where senior students and peer tutors share explanations, study strategies and tacit tips about how to approach demanding modules and assessment formats. Faculty mentors intervene mainly at key checkpoints (orientation events, feedback meetings, formal tutoring), while peer mentors and community administrators ensure day-to-day responsiveness, lowering the threshold for asking questions and normalising the sharing of difficulties and solutions.

A second set of knowledge flows concerns participation in extracurricular and research-oriented initiatives, where the mentoring ecosystem acts as an amplifier of opportunities and capabilities. For events such as the annual ERN in Bari, the core and cohort communities are used to identify interested students, form teams and circulate information on deadlines, logistics and expected contributions, while experienced participants mentor newcomers on how to design engaging demos, communicate research topics and position LUM Engineering alongside other universities. Similarly, the Team Ingenium initiative – an informal research and entrepreneurship-oriented group that, for the first time, was invited to present with a dedicated stand – illustrates how mentoring relations, community channels and faculty support can converge to transform a loosely connected group of students into a visible actor in regional innovation events, with spillover effects on skills, confidence and external recognition. In both cases, the ecosystem enables students to access not only information about opportunities, but also situated know-how on "how to be there": how to prepare, how to present themselves and how to learn from comparison with peers from other institutions.

The Ingenium initiative further illustrates how the mentoring ecosystem can translate into recognisable external outcomes. One of them, in December 2025, during the "Women in Big Data" hackathon hosted by the Innovation Department of the University of Salento, the team – combining female and male students from LUM

Engineering with a computer science student from another local university – received a Collaboration Award for its cross-institutional and gender-diverse composition and way of working.

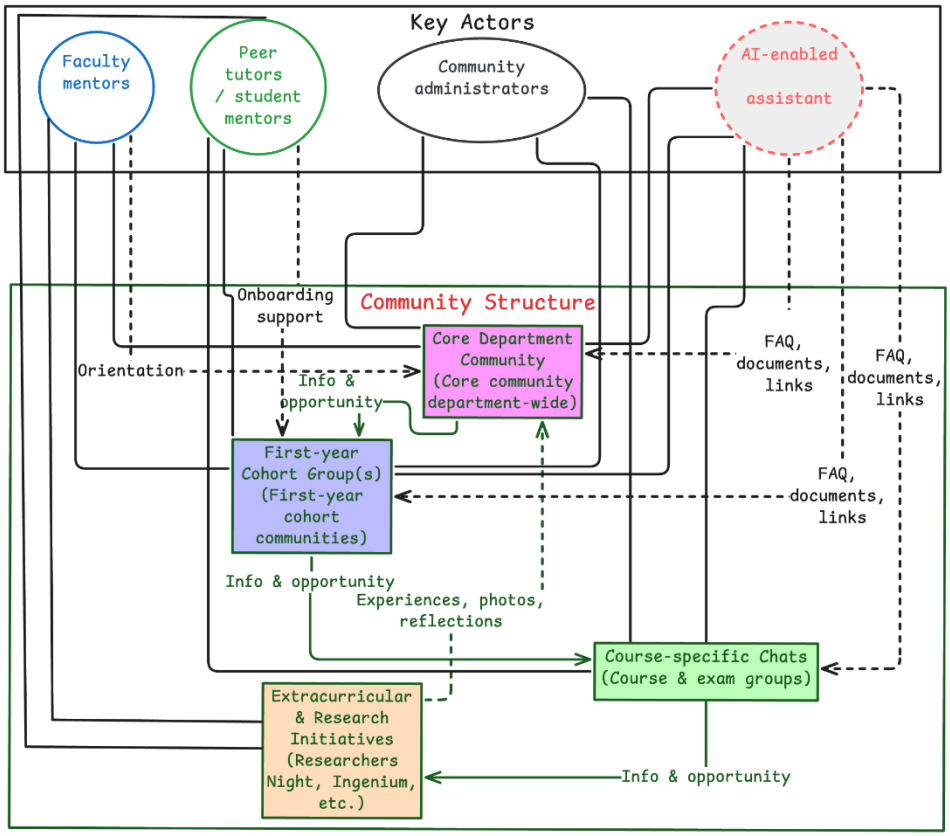


Figure 2: Knowledge flows in the "continuous mentoring" ecosystem (first-year support and extracurricular participation)

4.2 Key outcomes of the continuous mentoring ecosystem at LUM Engineering Department

The basic analytics of the digital communities provide additional insight into how the mentoring ecosystem operates in practice. Across several cohorts, first-year groups maintain participation rates that often exceed 70 – 80 % of enrolled students over multiple semesters, with cumulative message volumes in the thousands. This sustained activity suggests that the communities are not merely activated around onboarding events, but continue to function as day-to-day infrastructures for sharing information, study strategies and tacit tips about demanding courses and assessment formats. Figure 3 shows the growth of the core departmental noticeboard community over time, with the number of members increasing and then stabilising as new cohorts are added, confirming its role as a persistent backbone that consolidates communication across generations of students.

The temporal distribution of messages further reveals clear peaks in activity around key transition points, such as the start of the academic year, exam sessions and major initiatives like the annual ERN or entrepreneurship-oriented events.Project-based communities, such as those created for the Ingenium team and time-management workshops, show smaller membership but very high levels of active participation, reflecting their function as intensive learn- ing spaces where mentoring takes the form of close guidance and collaborative problem-solving. Figure 4 illustrates the evolution of total and active members in a first-year cohort community. Over multiple semesters, active participants consistently represent a large share of enrolled students, suggesting that the group functions as a sustained "safety net" rather than a short-lived onboarding channel.

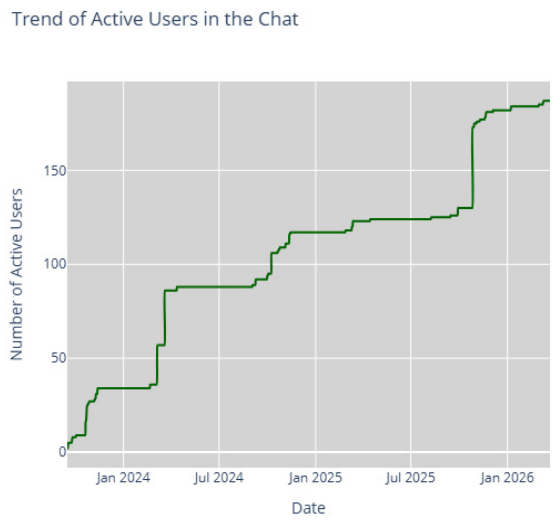


Figure 3: Growth of the core departmental noticeboard community (total members over time).

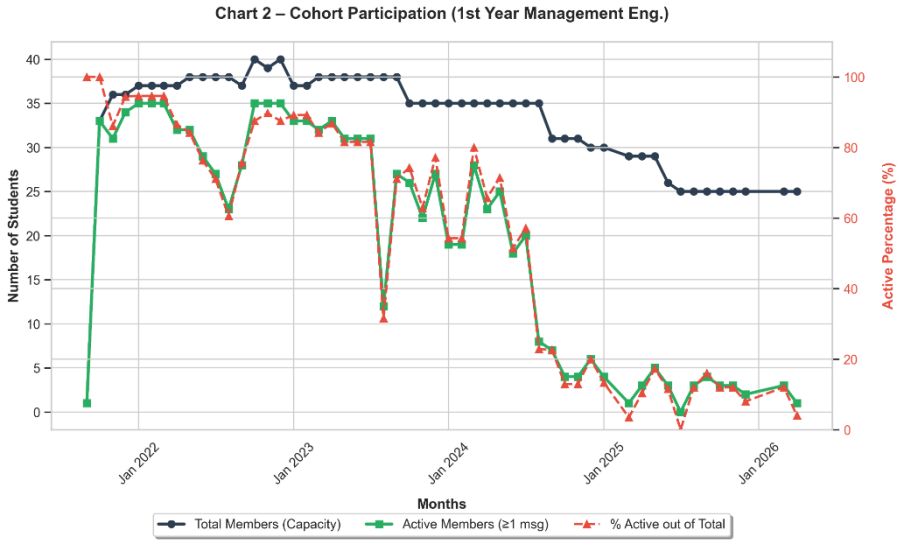


Figure 4: Total and active members in a first-year cohort community (students over time).

Figure 5 reports the monthly message volume in a project-oriented community. Activity peaks clearly align with key initiatives such as time-management workshops and entrepreneurship events, indicating that these smaller groups operate as intensive learning spaces where mentoring is tightly coupled with preparation, co-ordination and reflection around specific projects.

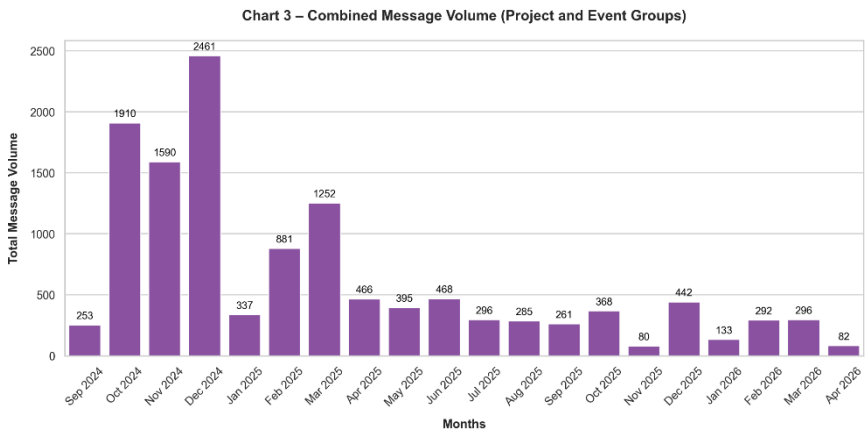


Figure 5: Monthly message volume in a project-oriented community within the mentoring ecosystem.

A preliminary semantic scan of messages across three representative communities – the core departmental noticeboard, a first-year cohort group and a project-oriented community – reinforces this interpretation. In the core community, the most frequent terms relate to schedules, deadlines, exams and institutional events, reflecting its role as a backbone for coordination and collective awareness. In the first-year cohort group, exam-related vocabulary and help-seeking expressions dominate, with recurrent references to “doubts”, “exams”, “notes” and “study groups”, indicating that students use this space as a low-threshold “safety net” to navigate early academic challenges and reduce the stigma of asking for support. In the project-oriented community, by contrast, the salience of terms such as “event”, “team”, “pitch” and “hackathon” points to a different function: these spaces concentrate the language of preparation, performance and external visibility, illustrating how the mentoring ecosystem also serves as a launchpad for participation in research and entrepreneurship initiatives beyond the classroom. Rather than offering a full semantic analysis, these patterns provide a first empirical glimpse into how different layers of the ecosystem host distinct but complementary knowledge and support flows.

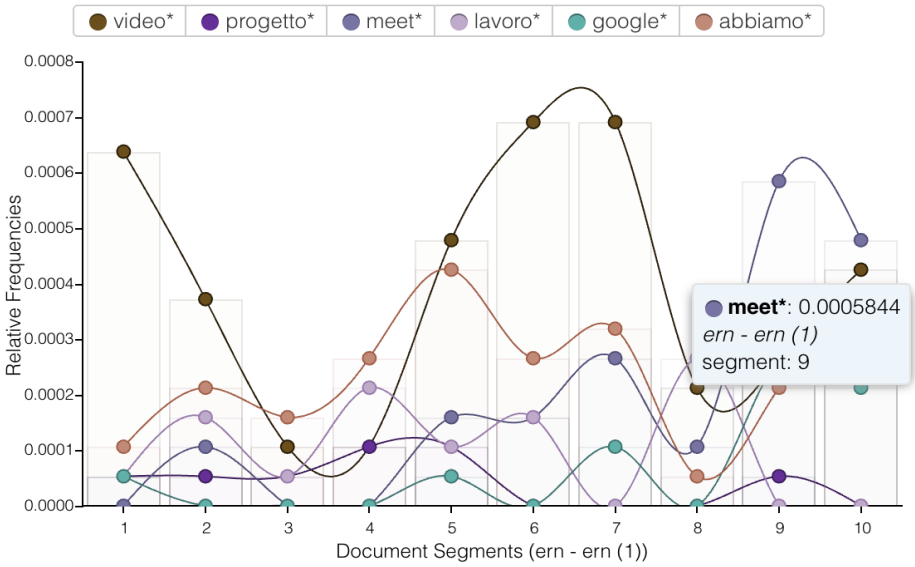


*Figure 6: Most frequent terms in mentoring-related community messages, highlighting dominant themes across coordination, help-seeking and project participation *(video,project,meet,job,google,we have)*

The two semantic views obtained with Voyant (FIgure 6 Cirrus and Figure 7 keyword trends) should be read as an exploratory “temperature check” of the mentoring ecosystem rather than as a full NLP analysis. The different community layers host distinct constellations of concerns and opportunities, opening a path for future work on systematic semantic monitoring of socio-digital mentoring ecosystems.

Figure 7: Word-cloud of the most frequent terms in mentoring-related community messages.

5 Theoretical Implications

The case extends mentoring research in at least three ways. First, it shows how formal, informal, peer and reverse-like mentoring practices can be configured not as separate interventions, but as layers of a coherent eco-system anchored in socio-digital infrastructures. While prior studies typically examine individual programmes or dyadic relationships in isolation, the LUM Engineering case illustrates how mentoring roles, community practices and digital tools can be orchestrated into a *continuous knowledge infrastructure* that spans cohorts, and transition points. This ecosystem perspective contributes to ongoing debates on how mentoring supports knowledge creation and sharing by highlighting the importance of persistent backbone communities and flexible satellite groups as complementary loci of learning.

Second, by combining communities-of-practice thinking with a multi-layered ecosystem lens, the study re-frames mentoring communities not only as informal networks emerging around shared concerns, but as deliberately cultivated socio-technical spaces where knowledge flows are shaped through light governance, role design and basic analytics. The growth and sustained activity of the core and cohort communities, together with the intensive exchanges in project-based groups, provide empirical grounding for the idea that digital communities can function as meso-level knowledge infrastructures that connect individual mentoring relationships with institutional strategies for student engagement and retention. This contributes to broader work on digital and entrepreneurial ecosystems by relocating the focus from firm–region dynamics to the micro-level of a university department.

Third, the preliminary use of activity metrics and semantic scans points to a promising avenue for integrating mentoring ecosystems with data-informed knowledge management. Rather than introducing sophisticated predictive models, the study suggests that simple indicators – such as participation rates, message volumes and dominant terms – can be mobilised as "temperature checks" of the mentoring ecosystem, helping educators monitor where support is flowing, where it may be lacking and how topics and concerns evolve over time. This opens up a research agenda on how AI-enabled tools and lightweight analytics can augment, rather than replace, human mentoring by providing new lenses for understanding and steering socio-digital knowledge infrastructures in engineering education and beyond.

6 Conclusions

This paper has argued that mentoring in an engineering education context can be fruitfully understood and designed as a sociodigital ecosystem that supports knowledge sharing, student success and participation in extracurricular initiatives. By documenting the *continuous mentoring* initiative at LUM's Engineering Department, the study has shown how a persistent core community, cohort-specific groups and project-oriented communities, supported by basic analytics and preliminary semantic scans, can be orchestrated into a coherent knowledge infrastructure. The case illustrates how this infrastructure lowers the threshold for asking questions, sustains everyday peer support and amplifies participation in opportunities such as research events and entrepreneurship competitions.

From a practical perspective, the paper has distilled three design principles for universities seeking to build mentoring-based knowledge infrastructures: anchor mentoring in a persistent core community that offers visibility and low-threshold access to help; cultivate flexible satellite communities around cohorts, courses and projects, where peer mentoring and experimentation can flourish; and augment, rather than replace, human mentoring with AI-enabled assistants and lightweight analytics for routine queries and ecosystem "health checks". Together, these principles highlight the importance of orchestrating people, practices and technologies as an integrated ecosystem rather than optimising isolated programmes.

Future research could extend this work in at least three directions. Longitudinal studies tracking student outcomes, such as retention, performance and early careers, would allow the effects of the mentoring ecosystem to be assessed more rigorously over time. Comparative analyses across institutions of different sizes, disciplines and national contexts could clarify which design principles are context-specific and which can be generalised beyond the LUM case. As AI-enabled tools and semantic analysis become more embedded in mentoring practices, further work should examine how chatbots, retrieval-augmented assistants and topic-sensitive dashboards reshape mentoring roles and knowledge codification, and how the exploratory semantic "temperature checks" used here can evolve into a more systematic method for monitoring topics, concerns and support needs across socio-digital communities

7 Ethics Declaration

This study draws on internal documents, project reports and anonymised digital traces from student communities within the Engineering Department. No sensitive personal data were collected, and no individual student is identifiable in the analysis. According to the current guidelines of the authors' institution, formal ethics approval was not required for studies based on organisational documents.

8 AI Declaration

AI tools (large language models) were used only at specific stages of the research process, namely to support the structuring and editing of the manuscript, to refine the wording of some sections (e.g. abstract, introduction and discussion) and to assist with the organisation and formatting of references. No AI tools were used for the conceptualisation of the research questions, research design, data collection, data preparation, data analysis or the generation of empirical findings. The authors reviewed and verified all AI-assisted outputs for accuracy and appropriateness, and cross-checked citations and bibliographic details against the original sources to avoid errors or misattribution.

References

AlAgha, I. (2014) 'An integrated approach to support knowledge representation, sharing and perception over Web 2.0', International Journal of Computer Applications, Vol. 97, No. 9, pp. 15–24.

Autio, E., Nambisan, S., Thomas, L.D.W. and Wright, M. (2018) 'Digital affordances, spatial affordances, and the genesis of entrepreneurial ecosystems', Strategic Entrepreneurship Journal, Vol. 12, No. 1, pp. 72–95.

Bjørnson, F.O. and Dingsøyr, T. (2005) 'A study of a mentoring program for knowledge transfer in a small software consultancy company', in Bomarius, F. and Komi-Sirviö, S. (eds) Product Focused Software Process Improvement (PROFES 2005), Lecture Notes in Computer Science, Vol. 3547, pp. 245–256, Springer, Berlin.

Bryant, S.E. (2006) 'The impact of peer mentoring on organizational knowledge creation and sharing: An empirical study in a software firm', Group & Organization Management, Vol. 31, No. 3, pp. 319–338.

Fowler, J.L., Fowler, D.S. and O'Gorman, J.G. (2019) 'Worth the investment? An examination of the organisational outcomes of a formal structured mentoring program', Australian Journal of Psychology, Vol. 71, No. 4, pp. 396–408.

Hall, M. (2003) Knowledge Management and the Codification of Knowledge in the UK Post Office, PhD thesis, The Open University, Milton Keynes.

Karkoulian, S., Halawi, L.A. and McCarthy, R.V. (2008) 'Knowledge management, formal and informal mentoring: An empirical investigation in Lebanese banks', The Learning Organization, Vol. 15, No. 5, pp. 409–420.

Romaldi, V. (2002) 'Collaborative technologies for knowledge management: Making the tacit explicit?', in Informing Science InSITE – Where Parallels Intersect, pp. 1358–1364.

Short, T. (2014) 'Workplace mentoring: An old idea with new meaning (Part 1)', Development and Learning in Organizations, Vol. 28, No. 1, pp. 8–11.

Short, T. (2014) 'Workplace mentoring: An old idea with new meaning (Part 2)', Development and Learning in Organizations, Vol. 28, No. 2, pp. 3–6.

Srivastava, S.B. (2015) 'Network intervention: Assessing the effects of formal mentoring on workplace networks', Social Forces, Vol. 94, No. 1, pp. 427–452.

Süveges, M. and Kurucz, A. (2024) 'Focus on onboarding process: Examining mentoring and training programs from the perspective of HR and employees', Journal of Infrastructure, Policy and Development, Vol. 8, No. 12, Article 8733.

Trunfio, D.A. (2024) Knowledge Management e l'AI Generativa – L.L.M. Base Solutions, Master's thesis, LUM Giuseppe Degennaro University, Casamassima.

Wenger, E. (1998) Communities of Practice: Learning, Meaning, and Identity, Cambridge University Press, Cambridge.

Yusof, A.N.M., Ahmad, N., Nirmala, N. and Lishudzaimah, N. (2012) 'Quality and effectiveness of knowledge management transfer using mentor–mentee program and on-the-job training in workplace', in Proceedings of the 2012 International Conference on Innovation Management and Technology Research (IMTIR 2012), IEEE.

AI Anxiety and Unemployment Anxiety: Moderating Role of Self-Efficacy in Cross-Cultural Study

Coşkun Akça[1], Albena Antonova[2], Anikó Csepregi[3], Zsuzsa Darida[3]
[1]Kastamonu University, Kastamonu, Turkey
[2]University of Sofia, Sofia, Bulgaria
[3]University of Pannonia, Veszprém, Hungary

Abstract

Artificial intelligence (AI) is rapidly transforming labour market structures, raising concerns not only about employment prospects but also about the broader psychological implications of technological change. Among university students, these concerns often emerge as heightened uncertainty regarding future employability. In this context, AI-related anxiety can be understood as a forward-looking stressor that shapes how individuals interpret knowledge about future work.

This study examines the relationship between AI anxiety and unemployment anxiety, with a particular focus on the moderating role of AI self-efficacy. A quantitative, cross-sectional design was applied using survey data collected from 510 university students in Turkey, Hungary, and Bulgaria. Established measurement scales were used to assess AI anxiety, AI self-efficacy, and unemployment anxiety. Data were analysed using SPSS and Hayes' PROCESS Macro (Model 1), alongside reliability, validity, and correlation analyses.

The results show that AI anxiety is positively associated with unemployment anxiety, suggesting that a large number of students tend to interpret AI-driven changes as potential threats to their future employment opportunities. Additionally, AI self-efficacy demonstrates a significant buffering effect: higher levels of perceived competence in using and understanding AI weaken the relationship between AI anxiety and unemployment anxiety. This points to the role of psychological resources in shaping how AI-related knowledge is cognitively and emotionally processed.

By integrating AI-related perceptions with perspectives on job insecurity, stress, and knowledge interpretation, the study contributes to the understanding of how technological change influences career-related anxiety. The findings suggest that AI anxiety should be conceptualised not merely as a technological attitude but as a mechanism through which individuals engage in sensemaking of AI-mediated knowledge about future employment.

From a practical perspective, the results underline the importance of strengthening AI-related competencies within higher education. Beyond technical skills, targeted interventions that enhance students' sense of control and interpretive capacity may play a key role in mitigating employment-related uncertainty in AI-driven labour markets.

Keywords: AI anxiety, AI self-efficacy, unemployment anxiety, university students

Paper type: Academic Research Paper

1 Introduction

In recent years, artificial intelligence (AI) has emerged not only as a technological innovation but also as a knowledge system that shapes the creation, interpretation, and use of knowledge (Hadjimichael et al., 2023; Mora et al., 2021). AI is recognised as a General-purpose technology with substantial economic and social effects (Crafts, 2021, Jain, 2025). Higher education provides a suitable context to examine this process, as it functions both as a knowledge ecosystem and as a setting for preparing students for the future labour market (Pandya and Wang, 2024). In this context, the growing diffusion of AI not only transforms modes of work but also influences how individuals make sense of and evaluate their future employment opportunities (Jin et al., 2024; Kim and Lee, 2024; Türk et al., 2025).

A stream of the literature examines the effects of AI mainly from automation, efficiency, or labour market displacement perspectives, emphasising the structural changes it triggers and their consequences (Erebak and Turgut, 2021; Chung et al., 2025). However, fewer studies examine how individuals make sense of AI-mediated knowledge, although organizational and cognitive research highlights the central role of sensemaking processes (Sandberg and Tsoukas, 2020; Whittle et al., 2023). A growing body of research shows that AI is not merely a source of information but a mediating system that shapes the structuring and sensemaking of knowledge (Jin et al., 2024; Kim and Lee, 2024; Li et al., 2025; Liu and Ye, 2026). This perspective reflects a shift from technological determinism to an interpretive view that foregrounds individual cognitive processing (Sandberg and Tsoukas, 2020).

Individual cognitive sensemaking processes are shaped by various subjective factors; therefore, uncertainty cannot be seen as determined solely by objective external conditions. Labour market uncertainty instead emerges as the outcome of individuals' sensemaking of knowledge, in which future-related information,

perceptions, and expectations are interconnected in complex ways (Shoss, 2017; Ganson et al., 2021; Chung et al., 2025). Prior psychological research shows that the perception of uncertainty is closely linked to cognitive processing and emotional responses, particularly anxiety, which can distort how information is interpreted and how future outcomes are inferred (Lamba et al., 2020; Sandhu et al., 2023). Several empirical studies confirm the relationship between AI-related anxiety and employment uncertainty (Erebak and Turgut, 2021; Uçar et al., 2025; Yaşar and Karagucuk, 2025), yet they do not examine in sufficient depth the underlying cognitive and psychological mechanisms behind this relationship.

Mechanism-based models that integrate knowledge management and human–AI cognition perspectives and explain how individuals make sense of AI-mediated knowledge, remain limited. Existing research highlights the role of self-efficacy as a cognitive resource and coping mechanism (Bandura, 2012; Peltier et al., 2022; Cabrera-Aguilar et al., 2023) and shows that AI-related competencies and attitudes influence the perception of anxiety (Kim and Lee, 2024; Li et al., 2025; Ren et al., 2025). However, these approaches rarely integrate sensemaking processes of AI-mediated knowledge and the emergence of uncertainty within a coherent theoretical framework.

This study addresses this gap by conceptualising AI as a knowledge-mediating system through which individuals make sense of their future employment opportunities, where cognitive and emotional mechanisms play a key role in the emergence of uncertainty. The study aims to examine how AI-related anxiety contributes to employment uncertainty and how AI-related self-efficacy moderates this relationship. The theoretical contribution lies in identifying the psychological mechanisms linking knowledge interpretation and the perception of uncertainty, and in integrating knowledge management and human–AI interaction perspectives within a coherent, mechanism-based model.

2 Theoretical Background

2.1 AI as a knowledge-based interpretive system

In the literature, artificial intelligence is traditionally viewed as a technology that automates tasks and improves efficiency (e.g., Bhargava et al., 2020; Pandya and Wang, 2024). However, a growing body of research shows that AI plays a more complex role in knowledge creation and interpretation (Hadjimichael et al., 2023; Sandberg and Tsoukas, 2020; Mora et al., 2021; Chester and Allenby, 2023). AI not only provides information but also mediates the creation, filtering, and sensemaking of knowledge, thereby shaping which information becomes accessible and how it is presented to individuals (Jin et al., 2024; Türk et al., 2025; Chester and Allenby, 2023). Consequently, AI can be understood as a knowledge-mediating infrastructure that shapes perception and decision-making (Chester and Allenby, 2023).

From a knowledge management perspective, knowledge is not merely objective information but a context-dependent phenomenon that emerges through sensemaking (Hadjimichael et al., 2023; Sandberg and Tsoukas, 2015). This view is particularly relevant for AI, where meaning-making occurs within sociomaterial interactions between technological systems and users (Mora et al., 2021). In this sense, AI is not simply a tool but functions as an extension of cognitive processes, shaping how individuals make sense of future labour market opportunities and risks (Chiriatti et al., 2025). Accordingly, AI can be understood as a knowledge-based interpretive system through which individuals interpret and structure future-oriented information, thereby shaping their perceptions of employment realities.

2.2 Psychological mechanisms of AI-mediated knowledge interpretation and human–AI cognition

The sensemaking of AI-mediated knowledge is not merely an information-processing activity but the result of complex cognitive and emotional mechanisms that shape how individuals respond to technological change (Sandberg and Tsoukas, 2020; Whittle et al., 2023; Jin et al., 2024; Erebak and Turgut, 2021). Among these mechanisms, self-efficacy, anxiety, and trust play a central role, jointly shaping human–AI cognition. This also highlights that meaning making is not a neutral process but unfolds through cognitive and linguistic mechanisms that shape how individuals make sense of information (Whittle et al., 2023).

Self-efficacy is a key determinant of behaviour, shaping the extent to which individuals feel capable of handling a given situation (Bandura, 2012). In the context of AI, it reflects a sense of competence and control in using the technology, enabling more structured and adaptive interpretation of knowledge (Kim and Lee, 2024; Ren et al., 2025). Prior empirical studies show that higher self-efficacy reduces anxiety and supports coping strategies, thereby stabilising the processing of AI-mediated information (Peltier et al., 2022; Cabrera-Aguilar et al., 2023; Li et al., 2025).

In contrast, anxiety distorts cognitive processes, particularly in uncertain situations. Research shows that anxiety impedes adaptive learning and accurate interpretation of information, leading to increased perceived uncertainty (Lamba et al., 2020; Sandhu et al., 2023). AI-related anxiety is especially important, as the rapid pace and unpredictability of technological change may be perceived as a threat, increasing employment uncertainty (Erebak and Turgut, 2021; Uçar et al., 2025; Yaşar and Karagucuk, 2025).

Trust also plays a key role in the sensemaking of AI-mediated knowledge, as it shapes the extent to which individuals rely on information generated by the technology. Trust is not merely a technical issue but a cognitive and social construct that emerges through interactions among the individual, the system, and the context (Afroogh et al., 2024; Li et al., 2024). An appropriate level of trust facilitates information acceptance and reduces perceived risk, whereas both excessive and insufficient trust may distort knowledge interpretation (Jiang et al., 2025).

These mechanisms jointly shape how individuals process AI-mediated knowledge and interpret its labour market implications. They can be understood as micro-level cognitive and emotional mechanisms that shape perceptions and decisions regarding future employment.

2.3 Career-related uncertainty as a knowledge-based construct

In traditional approaches, uncertainty is often viewed as a state determined by objective external factors; however, a growing body of research emphasises its cognitive and interpretive nature (Sandberg and Tsoukas, 2020; Hadjimichael et al., 2023; Whittle et al., 2023; Ferguson et al., 2024). Labour market uncertainty is not only a consequence of structural changes but also the result of how individuals make sense of information and expectations (Shoss, 2017; Chung et al., 2025).

Psychological research shows that the perception of uncertainty is closely linked to how individuals process available information and the meanings they assign to it (Ganson et al., 2021). In digital environments, this process is particularly complex, as uncertainty evolves dynamically through the integration of multiple information sources and signals (Ferguson et al., 2024; Bhargava et al., 2020). In addition, individual differences in information processing and tolerance for uncertainty may lead to substantial variation in perceptions (Sandhu et al., 2023).

In the context of AI, this means that the same technological environment may be interpreted differently by individuals, leading to varying levels of employment uncertainty (Bhargava et al., 2020; Pandya and Wang, 2024). Uncertainty is therefore not a direct outcome of AI's objective effects but rather the result of how individuals make sense of AI-mediated knowledge and future-oriented information.

2.4 Conceptual model and hypotheses

Based on the above theoretical considerations, this study proposes a mechanism-based model in which AI-related anxiety has a direct effect on employment uncertainty, while self-efficacy moderates this relationship. The model assumes that AI does not directly shape uncertainty but exerts its effects through sensemaking processes of AI-mediated knowledge.

AI-related anxiety may distort information processing and increase negative expectations about the future (Erebak and Turgut, 2021; Uçar et al., 2025), whereas self-efficacy can structure and stabilise these sensemaking processes (Bandura, 2012; Kim and Lee, 2024). Accordingly, self-efficacy influences not the effects of the technology itself but their interpretation, thereby playing a key role in the emergence of uncertainty.

Although the empirical model examines direct relationships, these reflect underlying cognitive and sensemaking processes that shape how AI-mediated knowledge is processed at a theoretical level. Accordingly, the model not only describes relationships between variables but also shows how technological change translates into individual-level uncertainty through sensemaking processes.

Based on these theoretical considerations, the following hypotheses are proposed:

- *H1: AI-related anxiety is positively associated with unemployment anxiety, as it may distort the sensemaking of AI-mediated information and reinforce uncertainty about the future.*
- *H2: AI-related self-efficacy weakens the relationship between AI-related anxiety and unemployment anxiety, as it enables more structured and controlled sensemaking of AI-mediated knowledge.*

The proposed model provides a framework for interpreting these relationships and examines how psychological responses to technological change translate into uncertainty about future employment (Figure 1). The model assumes that uncertainty does not arise directly from the effects of technology but from how individuals engage in sensemaking of AI-mediated knowledge, a process shaped by individual cognitive and psychological resources.

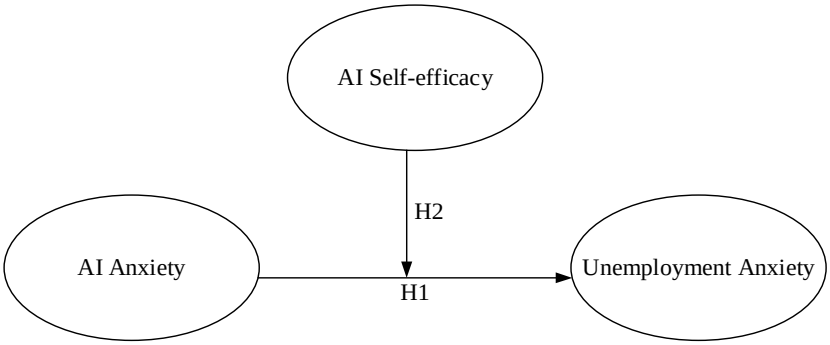


Figure 1: Interpretation-based research model

The model empirically represents direct relationships; however, at the theoretical level, these effects can be understood through cognitive and emotional mechanisms operating within the sensemaking of AI-mediated knowledge.

3 Methodology

This study used a cross-sectional survey design to examine the relationship between AI anxiety and unemployment anxiety among university students, with AI self-efficacy specified as a moderator. The data were collected in three countries (Turkey, Hungary, and Bulgaria), allowing the proposed relationships to be examined in a comparative higher education context rather than within a single national setting.

3.1 Research Sample

The sample consisted of 510 university students, with 170 respondents from each country. Data were collected through an anonymous online survey, and to ensure comparability across the three national subsamples, a balanced sampling approach was applied: from the initial 550 responses (Türkiye 180, Hungary 200, Bulgaria 170), 170 cases were retained from each country for the final analysis to ensure comparability across countries and to avoid potential bias resulting from unequal group sizes.

Participation was voluntary, and respondents were informed about the aim of the study before completing the questionnaire. Demographic characteristics are presented in Table 1. The final sample was predominantly female (n = 309), and the largest age group was 21–23 years (n = 265), which is consistent with the expected profile of the target population.

This design allows for a controlled comparison of AI-related perceptions across different higher education contexts.

Table 1**:** Demographic characteristics

Age			**Country**			**Total**
			Turkey	**Hungary**	**Bulgaria**	
18-20	Gender	Male	49	10	18	77
		Female	65	20	24	109
	Total		114	30	42	186
21-23	Gender	Male	22	36	37	95
		Female	24	84	62	170
	Total		46	120	99	265
24-26	Gender	Male	2	8	5	15
		Female	4	6	10	20
	Total		6	14	15	35
27+	Gender	Male	2	3	9	14
		Female	2	3	5	10
	Total		4	6	14	24
Total	Gender	Male	75	57	69	201
		Female	95	113	101	309
	Total		170	170	170	510

3.2 Measures

All items were assessed on a five-point Likert-type scale, and the questionnaire also included basic demographic variables such as gender, age, and country of residence. Unemployment anxiety was measured with the scale developed by Şanli et al. (2023), AI self-efficacy with the instrument proposed by Türk et al. (2025), and AI anxiety with the short-form measure developed by Batuk et al. (2025). In the case of the latter, only the AI anxiety dimension was used, as it was directly aligned with the focus of the present study.

Detailed psychometric properties of the original scales are not reproduced here; instead, the focus is on their performance in the current dataset.

All scales demonstrated satisfactory reliability in the current sample.

3.3 Analytical approach

All analyses were conducted using SPSS. The analysis proceeded in three steps. First, the measurement properties of the scales were assessed through exploratory factor analysis, reliability testing, and validity checks. Second, descriptive statistics and correlations were examined to identify the basic pattern of associations among the focal variables. Third, the hypotheses were tested using Hayes' PROCESS Macro (Model 1) to examine the moderating effect of AI self-efficacy on the relationship between AI anxiety and unemployment anxiety.

4 Results

4.1 Measurement validation

The measurement model showed satisfactory psychometric performance across all three constructs (see Table 2). Reliability and validity indicators supported the use of the scales in the subsequent analyses, confirming that AI anxiety, AI self-efficacy, and unemployment anxiety were empirically distinct constructs.

Table 2**:** Reliability and Validity Values

Variable	Factor Loading	Total explained variance	KMO	Bartlett's	Sig.	Alfa	AVE	CR
AI anxiety	798-882	71,625	,688	505,538	,000	,802	0.71	0.88
AI self-efficacy	767-887	68,256	,864	1356,309	,000	,880	0.68	0.91
Unemployment anxiety	539-863	56,177	,930	2834,822	,000	,910	0.56	0.92

4.2 Main relationships

The correlation analysis revealed a clear substantive pattern(see Table 3). AI anxiety was positively associated with unemployment anxiety, indicating that stronger concerns about AI tend to be accompanied by stronger concerns about future employment. By contrast, AI self-efficacy was negatively related to unemployment anxiety, suggesting that students with higher AI-related competence report lower levels of employment-related concern.

At the same time, AI anxiety and AI self-efficacy were not significantly related to each other, which implies that feeling anxious about AI is not simply the inverse of feeling competent in relation to it. This distinction is theoretically important because it suggests that the two constructs capture different psychological responses to the same technological environment.

Taken together, these findings move beyond a simple technology-effect explanation and indicate that unemployment-related concern is shaped both by threat-oriented perceptions of AI and by individuals' perceived capacity to cope with AI-related change.

Table 3**:** Descriptive Statistics, Correlation Matrix, and Discriminant Validity

Variable	N	Mean	SD	Skewness	Kurtosis	1	2	3
AI anxiety	510	2,9477	,87726	-,002	-,278	0.843		
AI self-efficacy	510	3,9027	,72059	-,598	1,103	,059	0.825	
Unemployment anxiety	510	3,1765	,90627	-,183	-,505	,176**	-,124**	0.748

4.3 Moderation effects

The moderation analysis supports the proposed relationships. AI anxiety had a significant positive effect on unemployment anxiety, while AI self-efficacy showed a significant negative direct effect. Importantly, the

interaction term was also significant, indicating that the relationship between AI anxiety and unemployment anxiety varies across levels of AI self-efficacy.

The pattern of conditional effects suggests a buffering mechanism. Among students with low AI self-efficacy, AI anxiety was strongly associated with unemployment anxiety. This relationship remained significant at average levels of self-efficacy, although with reduced strength. In contrast, at high levels of AI self-efficacy, the effect of AI anxiety was no longer statistically significant.

These results indicate that the impact of AI anxiety on employment-related uncertainty is contingent on individuals' psychological resources, particularly their perceived capability to interpret and manage AI-related information.

Table 4 presents the results of the moderation analysis.

Table 4**:** Hypothesis results

OUTCOME VARIABLE: Unemployment Anxiety						
Model Summary						
R	**R-sq**	**MSE**	**F**	**df1**	**df2**	**p**
0,281	0,079	0,761	14,459	3,000	506,000	,000
Model	**B**	**SE**	**t**	**p**	**LLCI**	**ULCI**
constant	3,184	0,039	82,332	,000	3,108	3,26
AI Anxiety	0,220	0,045	4,922	,000	0,132	0,308
AI Self-Efficacy	-0,207	0,055	-3,795	,000	-0,314	-0,100
Int_1	-0,200	0,050	-4,040	,000	-0,297	-0,103
Test of highest order unconditional interaction						
	R2-chng	**F**	**df1**	**df2**	**p**	
X*W	0,030	16,322	1,000	506,000	,000	
Conditional effects of the focal predictor at values of the moderator						
AI Self-Efficacy	**B**	**SE**	**t**	**p**	**LLCI**	**ULCI**
Low	0,365	0,062	5,905	,000	0,243	0,486
Mean	0,220	0,045	4,922	,000	0,132	0,308
High	0,076	0,052	1,456	0,146	-0,027	0,179

Int_1: AI Anxiety X AI Self-Efficacy, **X**= AI Anxiety, **W**=AI Self-Efficacy, **SE** = Standard error.

The results consistently show that the effect of AI-related anxiety on perceived employment uncertainty is not uniform but depends on the psychological resources shaping how individuals interpret technological change, particularly AI self-efficacy.

5 Discussion

The findings confirm that the relationship between AI-related anxiety and employment uncertainty is not a direct effect but emerges through individuals' cognitive and sensemaking processes. This aligns with approaches suggesting that uncertainty does not arise directly from the external environment but is constructed through interpretive processes (Sandberg and Tsoukas, 2020). In this sense, AI is not merely a technological tool but a knowledge-mediating system that shapes perception and meaning-making (Jin et al., 2024; Kim and Lee, 2024).

Psychological mechanisms also play a key role. The positive effect of AI-related anxiety supports the view that anxiety distorts information processing and inferences about the future (Lamba et al., 2020; Sandhu et al., 2023). In contrast, self-efficacy not only directly reduces uncertainty but, through its moderating role, shapes how individuals make sense of AI-related information, consistent with Albert Bandura's (2012) theory and related empirical findings (Peltier et al., 2022; Ren et al., 2025).

One of the most important findings is that at higher levels of AI self-efficacy, the effect of AI anxiety disappears, suggesting that individuals interpret technological change differently. This confirms that uncertainty does not arise directly from technological change but is shaped through sensemaking (Sandberg and Tsoukas, 2020), a process driven by cognitive and contextual sensemaking mechanisms (Whittle et al., 2023). This goes beyond the traditional "technology → effect" logic and highlights the interpretive nature of uncertainty (Ferguson et al., 2024).

The paper contributes to the knowledge management literature by showing that knowledge is not merely information but acquires meaning through interpretation. In this sense, AI is not only a source of information but a knowledge system in which knowledge is context-dependent and emerges through sensemaking (Hadjimichael et al., 2023). Moreover, AI functions as a sociomaterial environment that shapes sensemaking

through the interaction of technological and social factors (Mora et al., 2021). The effects of AI-mediated knowledge are therefore not deterministic but depend on individual sensemaking processes.

For higher education institutions, the findings suggest that AI system design should not be treated as a purely technical decision but as a process that shapes the visibility and sensemaking of knowledge.

In addition, higher education institutions can contribute to reducing students' employment-related anxiety by fostering AI-related self-efficacy, thereby supporting a more confident transition into the labour market.

This may be supported through a combination of technical training and soft skill development initiatives, including communication, self-awareness, and career preparation programmes.

It is interesting to highlight that the relatively high levels of AI self-efficacy, observed among students, do not necessarily correlate with their actual competencies in using AI. A substantial number of students claim that had not taken any formal course of training on AI, but at the same time, they report to feel as confident and very confident AI users. Generative AI instruments have low learning curve, thereby reinforcing users' perceived capability. In line with social cognitive theory (Bandura, 2012), the positive experiences of students with AI tools are likely to strengthen self-efficacy beliefs even if they lack deep conceptual understanding of AI technologies. This distinction is important, as even the operational level self-efficacy may still function as a psychological buffer against anxiety, as evidenced in the findings, while not necessarily indicating readiness or preparation for more complex or high-stakes AI applications.

6 Conclusion

This study contributes to the literature by examining the effects of AI not as direct technological outcomes but through sensemaking of AI-mediated knowledge. The findings suggest that employment uncertainty arises not only from technological change but from how individuals engage in sensemaking of AI-mediated knowledge.

From a theoretical perspective, the study integrates knowledge management and human–AI cognition perspectives, showing that AI functions as an interpretive system that shapes perceptions of future employment (Hadjimichael et al., 2023; Sandberg and Tsoukas, 2020). The mechanism-based approach explains how psychological factors mediate the effects of technological change through sensemaking processes (Whittle et al., 2023).

From a practical perspective, the findings highlight the importance of developing AI-related self-efficacy (Bandura, 2012). Higher education institutions should focus not only on technical skills but also on supporting students' sensemaking capabilities and cognitive readiness in AI-based environments. The results help organizations and educational institutions address the psychological dimensions of AI-related uncertainty.

A limitation of the study is its cross-sectional design, which does not allow clear causal inference. Future research could use longitudinal and qualitative approaches to better capture the dynamics of sensemaking processes and the role of knowledge interpretation in the emergence of uncertainty.

7 Ethics declaration

The research reported in this paper involved human participants and required ethical approval. Ethical clearance was obtained from the University of Pannonia Research Ethics Committee, on 04/02/2026, under approval number KEB 2/2026 (02.04.) Participation in the study was voluntary, and respondents were informed about the purpose of the research prior to completing the questionnaire. All data were collected anonymously and handled in accordance with applicable data protection and ethical research standards.

8 AI declaration

AI-based tool (ChatGPT) was used during the development of this paper to support language refinement, structuring, and clarity of expression. The authors critically reviewed, edited, and validated all generated content. No AI tools were used for data analysis, interpretation of results, or generation of empirical findings. The authors take full responsibility for the content of the manuscript.

References

Afroogh, S., Akbari, A., Malone, E., Kargar, M. and Alambeigi, H. (2024). 'Trust in AI: progress, challenges, and future directions', *Humanities and Social Sciences Communications*, 11(1), https://doi.org/10.1057/s41599-024-04044-8

Bandura, A. (2012) 'On the functional properties of perceived self-efficacy revisited', Journal of Management, 38(1), pp. 9–44. https://doi.org/10.1177/0149206311410606

Batuk, B., Aktu, Y. and Türk, N. (2025) 'Examining the psychometric properties of the short form artificial intelligence acceptance scale', *Çukurova University Journal of Social Sciences Institute*, 34 (The Transformation of Civilization: Artificial Intelligence), pp. 438–451. https://doi.org/10.35379/cusosbil.1695975

Bhargava, A., Bester, M. and Bolton, L. (2020) 'Employees' Perceptions of the Implementation of Robotics, Artificial Intelligence, and Automation (RAIA) on Job Satisfaction, Job Security, and Employability', *Journal of Technology in Behavioral Science*, 6(1), 106 - 113. https://doi.org/10.1007/s41347-020-00153-8

Cabrera-Aguilar, E., Zevallos-Francia, M., Morales-García, M., Ramírez-Coronel, A.A., Morales-García, S.B., Sairitupa-Sanchez, L.Z. and Morales-García, W.C. (2023) 'Resilience and stress as predictors of work engagement: the mediating role of self-efficacy in nurses', *Frontiers in Psychiatry*, 14, 1202048. https://doi.org/10.3389/fpsyt.2023.1202048

Chester, M. and Allenby, B. (2023) 'Infrastructure and the cognitive ecosystem: an irrevocable transformation', *Environmental Research: Infrastructure and Sustainability*, *3*(3), https://doi.org/10.1088/2634-4505/aced1f

Chung, Y.W., Im, S., Kim, J.E. and Yun, J.K. (2025) 'Artificial intelligence awareness, career resilience, job insecurity and behavioural outcomes', *Australian Journal of Psychology*, 77(1), 2559910. https://doi.org/10.1080/00049530.2025.2559910

Crafts, N. (2021) 'Artificial intelligence as a general-purpose technology: An historical perspective', *Oxford Review of Economic Policy*, 37(3), pp. 521–536. https://doi.org/10.1093/oxrep/grab012

Erebak, S. and Turgut, T. (2021) 'Anxiety about the speed of technological development: effects on job insecurity, time estimation, and automation level preference', *The Journal of High Technology Management Research*, 32(2), 100419. https://doi.org/10.1016/j.hitech.2021.100419

Ferguson, A., Turner, G. and Orben, A. (2024) 'Social uncertainty in the digital world. Trends in Cognitive Sciences', 28(4), 286 - 289. https://doi.org/10.1016/j.tics.2024.02.005

Ganson, K.T., Tsai, A.C., Weiser, S.D., Benabou, S.E. and Nagata, J.M. (2021) 'Job insecurity and symptoms of anxiety and depression among US young adults during COVID-19', *Journal of Adolescent Health*, 68(1), pp. 53–56. https://doi.org/10.1016/j.jadohealth.2020.10.008

Hadjimichael, D., Pyrko, I. and Tsoukas, H. (2023) 'Beyond tacit knowledge: how Michael Polanyi's theory of knowledge illuminates theory development in organizational research', *Academy of Management Review*. *49*(4), 919-934 https://doi.org/10.5465/amr.2022.0289

Jain, S. (2025) 'Comparative analysis of AI adoption and economic impact vis-à-vis earlier general-purpose technologies', *Kashmir Journal of Social Sciences*, 13(1), pp. 101–116.

Jiang, W., Li, D. and Liu, C. (2025) 'Understanding dimensions of trust in AI through quantitative cognition: Implications for human-AI collaboration', *PLOS One*, 20(7), https://doi.org/10.1371/journal.pone.0326558

Jin, G., Jiang, J. and Liao, H. (2024) 'The work affective well-being under the impact of AI', *Scientific Reports*, 14(1), 25483. https://doi.org/10.1038/s41598-024-75113-w

Kim, B.J. and Lee, J. (2024) 'The mental health implications of artificial intelligence adoption: the crucial role of self-efficacy', *Humanities and Social Sciences Communications*, 11(1), 1561. https://doi.org/10.1057/s41599-024-04018-w

Lamba, A., Frank, M.J. and FeldmanHall, O. (2020) 'Anxiety impedes adaptive social learning under uncertainty', Psychological Science, 31(5), pp. 592–603. https://doi.org/10.1177/0956797620910993

Li, R., Ouyang, S. and Lin, J. (2025) 'Mediating effect of AI attitudes and AI literacy on the relationship between career self-efficacy and job-seeking anxiety', *BMC Psychology*, 13(1), 454. https://doi.org/10.1186/s40359-025-02757-2

Li, Y., Wu, B., Huang, Y. and Luan, S. (2024) 'Developing trustworthy artificial intelligence: insights from research on interpersonal, human-automation, and human-AI trust', *Frontiers in Psychology*, 15. https://doi.org/10.3389/fpsyg.2024.1382693

Liu, X. and Ye, Z. (2026) 'The forgotten middle: how moderate self-efficacy amplifies the threat of AI through job insecurity', *Frontiers in Psychology*, 16, 1734254. https://doi.org/10.3389/fpsyg.2025.1734254

Mora, L., Kummitha, R. and Esposito, G. (2021) 'Not everything is as it seems: Digital technology affordance, pandemic control, and the mediating role of sociomaterial arrangements', *Government Information Quarterly*, 38(4), 101599. https://doi.org/10.1016/j.giq.2021.101599

Pandya, S. and Wang, J. (2024), 'Artificial intelligence in career development: a scoping review', *Human Resource Development International*, 27(3), 324 - 344. https://doi.org/10.1080/13678868.2024.2336881

Peltier, J.W., Chennamaneni, P.R. and Barber, K.N. (2022) 'Student anxiety, preparation, and learning framework for responding to external crises: the moderating role of self-efficacy as a coping mechanism', *Journal of Marketing Education*, 44(2), pp. 149–165. https://doi.org/10.1177/02734753211036500

Ren, J., Guo, J. and Li, H. (2025) 'Linking digital competence, self-efficacy, and digital stress to perceived interactivity in AI-supported learning contexts', *Scientific Reports*, *15*(1), https://doi.org/10.1038/s41598-025-18873-3

Sandberg, J. and Tsoukas, H. (2020) 'Sensemaking Reconsidered: Towards a broader understanding through phenomenology', *Organization Theory*, 1(1), https://doi.org/10.1177/2631787719879937

Sandhu, T., Xiao, B. and Lawson, R. (2023) 'Transdiagnostic computations of uncertainty: towards a new lens on intolerance of uncertainty', *Neuroscience & Biobehavioral Reviews*, 148. https://doi.org/10.1016/j.neubiorev.2023.105123

Şanli, A.G., Yazan, Z.Ö., Kirazcı, F. and Büyükgöze-Kavas, A. (2023) 'Development of the Job Finding Anxiety Scale for University Students: validity and reliability study', *The International Journal of Economic and Social Research*, 19(2), pp. 349–364.

Shoss, M.K. (2017) 'Job insecurity: an integrative review and agenda for future research', *Journal of Management*, 43(6), pp. 1911–1939. https://doi.org/10.1177/0149206317691574

Türk, N., Batuk, B., Kaya, A. and Yıldırım, O. (2025) 'What makes university students accept generative artificial intelligence? A moderated mediation model', *BMC Psychology*, 13, 1257. https://doi.org/10.1186/s40359-025-03559-2

Uçar, M., Çapuk, H. and Yiğit, M.F. (2025) 'The relationship between artificial intelligence anxiety and unemployment anxiety among university students', *Work*, 80(2), pp. 701–710. https://doi.org/10.1177/10519815241290648

Whittle, A., Vaara, E. and Maitlis, S. (2023) 'The role of language in organizational sensemaking: An integrative theoretical framework and an agenda for future research', *Journal of Management*, 49(6), 1807–1840. https://doi.org/10.1177/01492063221147295

Yaşar, H. and Karagucuk, V. (2025) 'The effect of artificial intelligence anxiety on career decidedness among students in English-related departments at universities', *Discover Artificial Intelligence*, 5(1), 48. https://doi.org/10.1007/s44163-025-00284-y

MatchIS: A Digital Platform Designed to Address Critical Gaps in Industrial Symbiosis Marketplaces

Giulia Lippi[1], Luca Colomba[2], Paolo Garza[2], Enrica Vesce[1]
[1]Department of Management, University of Turin, Turin, Italy
[2]Department of Control and Computer Engineering, Polytechnic of Turin, Turin, Italy

Abstract

Industrial Symbiosis is widely recognised as a key strategy for fostering the ecological transition of production systems. This model allows companies to reduce their environmental impacts by exchanging unused resources, whilst maintaining their economic competitiveness and lowering their operational costs. Despite the promising outcomes, its adoption is still hindered by several barriers. Informational gaps related to the lack of data-sharing among companies, technological lock-ins, unclear legislation, and difficulties in assessing the economic viability of symbiotic exchanges are some of the obstacles that companies face when engaging in Industrial Symbiosis. Within this context, digitalisation and, in particular, Information and Communication Technologies, may help to widen symbiotic possibilities and connect companies across broader geographical boundaries. This is testified by the rise of digital platforms promoting Industrial Symbiosis, which can be classified into three main groups according to their core functionalities: good practice repositories, marketplaces, and training solutions. However, most Industrial Symbiosis platforms are little used and often tend to become static spaces, especially in the case of open and public platforms.

The aim of this paper is to present the design of the MatchIS platform, which aims to foster Industrial Symbiosis adoption and help companies address their main challenges. The design of MatchIS was developed through a literature review of IS barriers and semi-structured interviews with representatives of business associations and Chambers of Commerce, which provided both theoretical grounding and empirical insights into the limitations of existing platforms. Building on these results, MatchIS is designed as an open, public, and integrated platform that combines marketplace functionalities with a multidisciplinary support system. The platform enables firms to exchange information on production residues and valorisation technologies, while also providing access to legal, economic, environmental, and certification-related guidance. Distinctively, the platform explicitly incorporates multiple symbiotic roles — producers, utilisers, and transformers — thereby overcoming the binary perspective that characterises most existing solutions.

The paper contributes to the literature by proposing an integrated and accessible platform model that addresses multiple IS barriers simultaneously, demonstrating that advanced support services typically associated with private platforms can be delivered within an open and publicly accessible environment. However, this study is conceptual and design-oriented, focusing on how the identified barriers to IS have been translated into specific platform features. Empirical validation of the platform's effectiveness is left for future research.

Keywords: Industrial Symbiosis, Digital Platform, Marketplace, Waste, Byproducts

Paper type: Academic Research Paper

1 Introduction

The mounting environmental pressures associated with contemporary production and consumption systems render the transition to more sustainable economic models imperative. Institutions worldwide are advocating for new paradigms that can decouple economic growth from resource depletion and environmental degradation. Within this context, various production and consumption strategies and models oriented to sustainability have been consolidated under the overarching framework of the Circular Economy (Bocken et al, 2016). Indeed, this paradigm involves the sharing, leasing, reusing, repairing, refurbishing and recycling of existing materials and products to extend their lifespan. Circular strategies can be pursued and implemented at various levels, e.g., company, inter-firm, supply chain, municipal or regional scales. Among the various operational strategies underpinning the circular economy, Industrial Symbiosis (IS) has gained increasing attention as an effective approach to optimise resource use at the meso-level (Ghisellini, Cialani and Ulgiati, 2016), fostering collaboration among firms across different industrial sectors. The EU has recognised IS as a key circular operational tool to enable the ecological transition of manufacturing systems (European Commission, 2020). IS refers to a production strategy that connects different productive actors, belonging to different production segments, that establish physical synergies for the exchange and reuse of production residues and effluents (Lippi, Evola and Vesce, 2025). Therefore, IS enables circular integration across conventional supply chains by promoting resource optimisation before the consumption stage. Symbiotic exchanges are indeed limited to pre-consumer residues, such as by-products, waste, and end-of-waste materials arising from industrial

and manufacturing activities (Chertow and Ehrenfeld, 2012). In particular, IS allows companies to reduce their environmental burden whilst maintaining their economic competitiveness (Wadström, Johansson and Wallén, 2021). Companies engaging in symbiotic practices reduce their amount of waste or their reliance on virgin raw materials and, as a consequence, decrease the operational costs associated with input sourcing or waste management.

Symbiotic practices have been well documented for at least thirty years (Domenech et al., 2019; Neves et al., 2019), and some authors argue that they have always existed (Desrochers, 2004). Despite the recent acknowledgement of IS and the related diffusion, most companies, especially small and medium enterprises, struggle to adopt it (Akthar et al. 2022). Several barriers still hinder the diffusion of IS (Golev, Corder and Giurco, 2015). Among these, informational barriers play a central role (Lombardi, 2017). Firms often lack adequate knowledge of other companies' production processes, resulting in limited awareness of potential synergies involving the exchange of residual resources. This lack of visibility is further compounded by concerns related to confidentiality and the protection of sensitive information, which reduces firms' willingness to share data about their operations (Fraccascia and Yazan, 2018). Additionally, establishing symbiotic relationships requires a high level of coordination and trust among actors, which is often difficult to achieve, especially in fragmented industrial contexts (Velenturf and Jensen, 2015). Regulatory uncertainty, particularly regarding the legal classification of residues as waste or by-products, further complicates the implementation of IS initiatives (Lybæk, Christensen and Thomsen, 2021).

Within this context, digitalisation can play a crucial role in overcoming these barriers. Information and communication technologies (ICTs) offer new opportunities to facilitate data sharing, enhance transparency, and support the identification of potential symbiotic exchanges (Antikainen, Uusitalo, and Kivikytö-Reponen, 2018; Fraccascia and Yazan, 2018; Krom, Piscicelli, and Frenken, 2022). In particular, digital platforms can connect geographically dispersed actors, enabling the matching of supply and demand for residual resources and expanding the scope of IS networks beyond local industrial clusters. As a result, a growing number of digital solutions have been developed to support IS implementation. However, many existing IS platforms have struggled to achieve sustained engagement and tangible results (Maqbool, Mendez-Alva and van Eetvelde, 2018). Several initiatives remain underutilised or evolve into static databases with limited active exchanges among users. This gap between technological potential and practical effectiveness highlights critical shortcomings in current platform designs.

Against this backdrop, this paper presents the design of MatchIS, a digital marketplace aimed at facilitating IS among companies. Developed within the framework of the Digital and Sustainable North-West project funded by the Italian National Recovery and Resilience Plan, MatchIS has been conceived to address the main limitations of existing IS platforms by integrating a marketplace for resource exchange with dedicated helpdesks providing multidisciplinary support.

The remainder of the paper is structured as follows: the next section reviews the literature on IS barriers and existing digital platforms; Section 3 outlines the methodological approach adopted for the design of the platform; Sections 4 and 5 present and discuss the results. Particular attention is given to illustrating how the functionalities incorporated in MatchIS address the barriers identified in the literature.

2 Literature Review

2.1 Industrial Symbiosis

This section provides an overview of the IS concept and of the barriers that have to be overcome to operationalise it.

Adopting a material perspective, IS can be defined as the exchange of unused resources, such as byproduct, waste, residual energy and wastewater, among companies belonging to different production sectors (Chertow and Ehrenfeld, 2012; Lippi, Evola and Vesce, 2025). Therefore, the traditional symbiotic set-up foresees the material flow from one producer to one utiliser of the symbiotic resource. Those two are the main symbiotic roles, considering the binary synergy approach (Mendez-Alva et al. 2022). However, recent literature has acknowledged the necessity of intermediary actors that make the resource fit for further use (Patricio et al., 2022; Prosman and Wæhrens, 2019), thereby adding a new role within the symbiotic set-up: that one of the transformers. In order to activate IS material synergies, intermediate steps may be required to prepare or treat the residues (Chertow and Ehrenfeld, 2012). These steps may encompass upgrading the materials, cleaning, refurbishment and sometimes recycling. Transformation activities may be carried either by the producer or the utiliser, or by a third independent party. The intermediary actor is often referred to as upcycling tenant (Martin and Eklund, 2011; Oliveira Pavan et al., 2021), namely industrial plants taking and transforming wastes and by-

products from several other industries which otherwise would not be reusable, to create a new valuable intermediate product deployable by other productive processes.

Although IS implementation may be associated with a number of economic, environmental and social benefits, its establishment is challenged by a range of barriers including technological lock-ins, economic disincentives, informational gaps, regulatory hurdles, organisational, uncertainty and risk-related barriers (Domenech et al., 2019; Golev, Corder and Giurco, 2015; Neves et al., 2020). Barriers are perceived differently depending on the role and the sector of the company. Nonetheless, technological and economical obstacles are the most significant (Lombardi, 2017). Technological barriers refer to the technology readiness level (TRL) of the production process involving the use of productive residue. Existing technologies do not always allow for the activation of a byproduct synergy. Reuse, recovery, and recycling routes often face technological lock-ins, hindering the adoption of innovative solutions. Economic barriers mainly refer to timing mismatches between residue generation and demand and lack of economic incentives to support the secondary materials market. In fact, a significant economic disincentive stems from low raw material prices, which heavily influences the secondary market. Regarding risk-related barriers, the two mostly discussed refer to the possibility that long-term value chain partnerships may reduce companies' flexibility to respond to market changes and to the fact that increasing the inter-dependency among the industries can increase their vulnerability with respect to price and quantity variations (Herczeg, Akkerman and Hauschild, 2018). The regulatory dimension is of particular importance and has a twofold effect. Policy instruments can either facilitate IS formation or impose restrictions on industrial owners that hinder potential projects through policy fragmentation (Domenech, 2021; Lybæk, Christensen and Thomsen, 2021). Governments can play a key role by adopting programs like tax reductions or pollution control, directly or indirectly creating economic incentives for companies that incorporate IS in their production systems. At the same time, restrictive policies hinder the establishment of symbiotic exchanges by imposing strict requirements for residue quality. Likewise, a fragmented legislation may blur either the definition of IS or of different types of waste, imposing unclear administrative procedures. Table 1 provides a summary of the main barriers that should be addressed to promote IS. Particularly, the focus was set on those that can be tackled through digitalisation. As already mentioned, digitalisation can play a pivotal role in fostering data-sharing and support companies in evaluating their symbiotic possibilities.

Table 1. Barriers to the implementation of IS, according to the scientific literature.

Economic	Division of income and costs between organizations
	Different investment cycles
	Fluctuations in demand for a particular commodity or product
	Variation in cost of resources across regions
	Increase in operational costs (i.e. waste transportation and waste treatment costs)
Technical / Operational	Changing operations design
	Mismatch between demand and supply;
	vulnerabilities due to material fluctuations
	Variations in, or issues related to, waste quality
	Lack of technology and infrastructure readiness
Legislative	Restrictive environmental legislation
	Uncertain legislation or lack of guidance (on compliance criteria, best environmental practices, permit conditions, etc)
	Difficulties in obtaining approval for waste reuse projects from the regulatory authorities
Social/Informational	Social isolation between organizations
	lack of information about other companies' by-products and waste flows
	Lack of willingness to collaborate
	Lack of necessary information or data; lack of training for implementing IS
	Lack of awareness of IS concepts
	Lack of information sharing
Environmental	Unknown benefit

2.2 Existing Industrial Symbiosis platforms

The primary objective of this section is to provide a comprehensive overview of state-of-the-art digital platforms dedicated to IS. These platforms play a pivotal role in addressing the barriers outlined in the previous section, particularly those related to information asymmetries, partner identification, and coordination among actors. Across Europe, numerous initiatives—often supported by European Union funding—have contributed to

the development of IS digital platforms, characterised by diverse objectives, core functionalities, sectoral focus, and geographical scope.

Considering their functional configuration, three main core functionalities can be identified across existing IS platforms (van Capelleveen, Amrit, and Yazan, 2018; Krom, Piscicelli, and Frenken, 2022). The first concerns good practices repositories. Within these digital spaces, successful case studies and best practices related to IS and, more broadly, to the Circular Economy, are collected and disseminated. The primary role of this type of platform is to support knowledge sharing by providing companies with examples of waste reduction, by-product valorisation, and sustainable resource management practices. While valuable from an informational perspective, these platforms often remain descriptive in nature and do not directly facilitate operational exchanges among firms. The second functionality that can be identified is related to marketplaces and matching tools. Within these digital environments, companies can offer and acquire resources, materials, and by-products, while also identifying potential partners for symbiotic exchanges. Therefore, these platforms provide virtual spaces for trading residual resources, supported by features such as search tools, material classification systems, and matchmaking algorithms. By making both the supply and demand of residual resources visible, they directly address informational barriers and facilitate the identification of compatible exchanges. Depending on their level of sophistication, these platforms may also include additional services such as logistics coordination, preliminary feasibility assessments, or support in establishing business relationships. Finally, another recurring functionality consists of training programmes. Training platforms provide educational resources, guidelines, and capacity-building programmes related to IS. These platforms aim to enhance firms' knowledge and skills, supporting the adoption of circular practices through training materials, workshops, and digital learning tools.

Nonetheless, it is important to note that the above-mentioned functionalities are not mutually exclusive categories. Rather than representing distinct and isolated platform types, they reflect core functionalities that digital platforms may incorporate to varying degrees. In practice, many platforms combine multiple functionalities within a single digital environment, adopting a more integrated approach (Krom, Piscicelli, and Frenken, 2022).

In addition to functional classifications, IS platforms can also be distinguished based on their level of accessibility. Open platforms provide free and unrestricted access to information and services. Semi-open platforms require user registration to access advanced features or detailed information, thereby balancing openness with controlled data sharing. Finally, closed platforms are typically operated by private organisations or consultancy providers, offering specialised and often tailored services, such as the facilitation of symbiotic exchanges and the management of industrial networks. This classification partially overlaps with that based on ownership, which distinguishes between public and private platforms.

Another relevant dimension of analysis concerns the geographical scope of IS platforms. Some platforms are designed to operate within a clearly defined territorial context, such as eco-industrial parks or specific industrial districts, where physical proximity facilitates symbiotic exchanges. Others operate at a broader scale, including regional or national platforms, while some digital solutions do not impose explicit geographical constraints, potentially enabling interregional or even international exchanges.

In Italy, the platform landscape is particularly dynamic, reflecting the commitment to sustainability. A prominent example is the Italian Circular Economy Stakeholder Platform (ICESP), which primarily functions as a repository of best practices and a knowledge-sharing hub. Concurrently, several Chambers of Commerce have implemented platforms such as Elenco Sottoprodotti, aimed at supporting the exchange of unused resources among local companies. Alongside public initiatives, private actors also play a significant role in the Italian IS ecosystem. Companies such as Sfridoo and Circularity offer specialised digital services that combine multiple functionalities, including marketplace and matching features, consultancy, and training. These platforms contribute to bridging informational and organisational gaps by connecting firms with symbiotic opportunities, providing expert guidance, and supporting the practical implementation of circular strategies.

Overall, the coexistence of multiple functionalities within individual platforms, together with differences in accessibility, ownership, and geographical scope, highlights the heterogeneity of current IS digital solutions. Table 2 classifies some of the existing Italian platforms according to these dimensions. As mentioned in the introduction, most IS-related digital initiatives struggle to remain operative and to effectively promote the implementation of symbiotic practices. For instance, closed and private platforms limit access to data and make it difficult to assess their actual effectiveness in facilitating symbiotic exchanges. On the other hand, many of the public and open matchmaking tools have slowly become static repositories, given a decrease in participation.

Table 2. Characterisation of some Industrial Symbiosis platforms, including MatchIS

Name	Core Functionalities	Access	Ownership	Geographical scope
Elenco Sottoprodotti	Marketplace	Semi-open	Public	Regional
Symbiosis	Marketplace and matchmaking	Closed	Public	District
Sfridoo	Marketplace and Good Practice Repository	Closed	Private	National
Circularity	Marketplace, Training and Consultancy	Closed	Private	National
ICESP	Good Practice Repository	Open	Public	National
INSIGHT	Good Practice Repository and Training	Closed	Public	European
MATCHIS	Marketplace and Consultancy	Open	Public	Regional

3 Methodology

This study adopts a design-oriented research approach to develop the MatchIS platform. Particularly, the objective is to translate insights related to IS barriers and existing platforms' limitations into concrete design features. The methodological approach consists of three main phases, as illustrated in Figure 1. The first phase relies on a literature review to identify the most important barriers to IS implementation. Then, three semi-structured interviews were conducted with representatives of business associations and Chambers of Commerce. These actors were selected due to their strategic role in facilitating inter-firm collaboration and their direct knowledge of the challenges faced by companies in adopting IS practices. The interviews aimed to understand which of the identified barriers are most perceived by local companies and which are the most influential problems of existing IS platforms that have hindered their utilisation. These complementary sources enabled a comprehensive understanding of both theoretical and practical limitations affecting existing IS platforms. Furthermore, the insights drove the design requirements for MatchIS. Specific functionalities were developed to address the most important barriers. Therefore, the platform architecture adopts an integrated approach, combining more core functionalities.

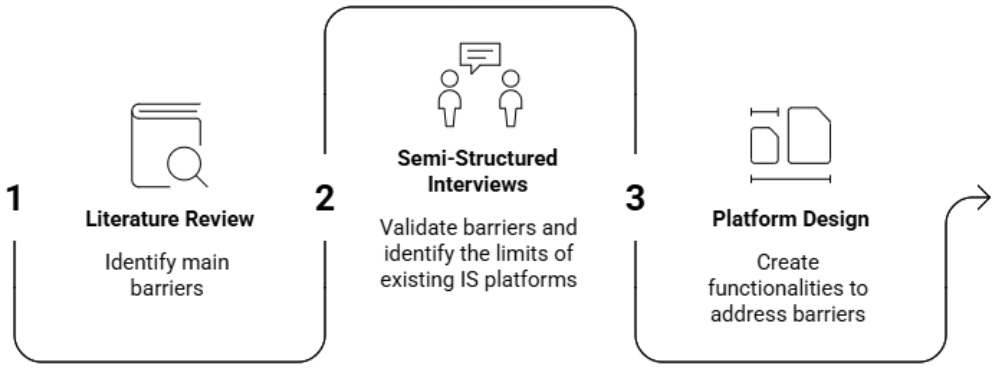


Figure 1. Methodological approach

4 Results

4.1 Existing Industrial Symbiosis platforms

The semi-structured interviews provided empirical support for several of the barriers identified in the literature. In particular, stakeholders consistently emphasised a limited awareness of IS among firms, especially small and medium-sized enterprises, which often lack a clear understanding of the practice as well as their potential symbiotic opportunities. This lack of awareness is closely linked to insufficient information on the availability and demand of specific resources, making it difficult for companies to identify suitable partners.

Regulatory uncertainty also emerged as a critical issue, particularly with regard to the distinction between waste and by-products. This ambiguity generates reluctance among firms to engage in symbiotic exchanges, especially within digital environments where concerns over legal compliance are amplified. Furthermore,

interviewees highlighted the difficulty in assessing the economic viability of IS practices, as companies often lack the necessary tools to evaluate potential cost savings.

Additional barriers include limited availability of technologies for waste valorisation, insufficient knowledge of possible valorisation pathways, and difficulties in quantifying and communicating environmental benefits. Overall, these findings confirm that IS implementation is hindered not only by informational gaps, but also by a broader lack of technical, economic, and regulatory support.

4.2 Critical gaps of existing platforms

The interviews also provided insights into the main limitations of existing IS platforms. First, good practice repositories, while useful for disseminating virtuous examples, often fail to provide sufficient detail to ensure replicability. As a result, they contribute only marginally to knowledge transfer and practical implementation.

Second, stakeholders expressed scepticism towards both open and closed marketplace platforms. On the one hand, companies are reluctant to share sensitive information on open platforms, particularly in the presence of regulatory uncertainty. On the other hand, closed platforms are often perceived as costly and restrictive, discouraging participation.

Third, most existing platforms adopt a binary approach to symbiotic exchanges, focusing exclusively on interactions between producers and utilisers. This perspective neglects the role of intermediary actors, thereby limiting the feasibility of more complex symbiotic configurations. Finally, only a limited number of platforms—typically private and closed—provide consultancy services or decision-support tools, which are, however, essential to assess the feasibility of IS initiatives.

In light of the above, the main objective is to design an open, public platform which combines a marketplace with a support system aimed at enhancing the feasibility and reliability of symbiotic transactions. In particular, the platform was designed to: (i) improve the visibility of supply and demand for production residues; (ii) explicitly include different symbiotic roles, such as producers, utilisers, and transformers; and (iii) provide access to specialised support services through a dedicated help-desk covering legal, economic, environmental, social, and certification aspects.

4.3 MatchIS design

This section presents the main outcomes of the research, namely the design and functional configuration of the MatchIS platform. As the platform has been recently deployed, the results are primarily conceptual and design-oriented, focusing on how the identified barriers to IS have been translated into specific platform features.

Building upon the findings illustrated in Section 4.1 and 4.2, MatchIS has been designed as an open, public, and integrated platform aimed at overcoming the identified barriers. Unlike many existing solutions, which either provide limited functionalities or operate as closed systems, MatchIS combines marketplace functionalities with a comprehensive support system, thereby adopting an integrated approach that is typically observed only in private platforms. This aspect represents a key distinguishing feature of MatchIS, as it delivers advanced services within an open and freely accessible environment.

The marketplace constitutes the core component of the platform and directly addresses informational barriers by enabling firms to publish structured advertisements related to both the supply and demand of production residues, as well as available valorisation technologies. The implementation of a standardised tagging system significantly enhances the visibility and comparability of information, allowing users to efficiently identify potential symbiotic opportunities based on material characteristics, quantities, and geographical proximity. The platform therefore incorporates filtering functionalities based on geographical and material criteria, and a semantic search function that allows users to navigate the marketplace efficiently.

Nonetheless, the key innovative feature of the MatchIS marketplace is the explicit inclusion of multiple symbiotic roles. In addition to producers and utilisers, the platform incorporates transformers, thereby overcoming the limitations of the traditional binary model. This enables the development of more complex and realistic symbiotic networks and facilitates the identification of valorisation pathways for materials that would otherwise remain unused. Not only producers and utilisers are allowed to publish advertisements, but also intermediary actors whose core business is the valorisation of waste. They can offer their technologies on the platform, enabling both knowledge transfer and the establishment of more sophisticated synergies. When publishing an advertisement, companies are required to specify the transformation needed and whether they perform such transformation and, if so, whether they are willing to act as intermediary for other companies.

Importantly, the platform separates the matching phase from the transactional phase, allowing companies to maintain control over negotiations and contractual agreements. This design choice contributes to reducing concerns related to confidentiality and trust, which emerged as critical barriers during the empirical analysis.

The marketplace is integrated with multidisciplinary helpdesks, which constitute a central element of the platform design. The helpdesks provide support across legal, economic, environmental, social, and certification domains. These services directly address several of the barriers identified in the literature and interviews. Within each helpdesk a list of frequently asked questions is presented, together with guidance on which public agencies can provide further consultancy. Therefore, these sections provide companies with structured guidance and practical information on key aspects of IS implementation. For instance, the legal helpdesk reduces regulatory uncertainty by clarifying the classification of production residues and the associated administrative procedures to be followed. Likewise, the certification helpdesk contains information on the schemes currently recognised by the European and Italian authorities to certify the engagement on IS, whilst the environmental helpdesk provides guidance on how to implement life cycle assessment analysis within the IS context. Of particular interest, within the economic helpdesk an automated mathematical modelling tool is proposed. The tool is based on linear programming and evaluates the costs that the utiliser has to bear in case of establishing a particular synergy. The model considers also logistics and transformation requirements. This tool enables firms to compare the economic convenience of using production residues instead of virgin raw materials, thereby supporting more informed decision-making.

Notably, these decision-support tools are freely accessible and designed to be user-friendly, lowering the entry barriers for companies with limited technical expertise. This represents a significant departure from existing platforms, where such services are typically restricted to private and fee-based solutions.

Overall, MatchIS represents an innovative and integrated approach to IS digitalisation, combining openness, accessibility, and advanced functionalities in a single platform. By simultaneously addressing informational, regulatory, economic, and organisational barriers, the platform contributes to enhancing both the feasibility and the scalability of IS practices.

5 Discussion

The results presented in the previous section provide several insights into the design of digital platforms for IS and their potential to address persistent implementation barriers. From a theoretical perspective, this study contributes to the literature by moving beyond the traditional classification of IS platforms based on isolated functionalities. Instead, this study supports the view that effective digital solutions should adopt a hybrid and integrated approach, combining multiple functionalities within a single platform. The design of MatchIS reflects this perspective, as it simultaneously incorporates marketplace, matching, and support-service features.

A second theoretical contribution concerns the explicit recognition of intermediary actors within IS systems. By introducing the role of transformers, the platform acknowledges that many symbiotic exchanges require additional processing steps before residues can be effectively reused. This expands the conventional binary perspective of IS interactions and provides a more realistic representation of industrial ecosystems.

Furthermore, the integration of multidisciplinary helpdesks highlights the importance of combining digital matching tools with expert knowledge. The findings suggest that the availability of information alone is not sufficient to enable IS adoption. Instead, firms require support in interpreting regulatory frameworks, assessing economic viability, and evaluating environmental implications. This reinforces the argument that digital platforms should not be conceived merely as information repositories or marketplaces, but as comprehensive service systems supporting decision-making processes. This is particularly relevant as integrated approaches combining digital marketplaces with consultancy services are more commonly associated with closed and private platforms, whereas MatchIS demonstrates the feasibility of delivering such functionalities within an open and publicly accessible system.

From a practical standpoint, the MatchIS platform offers several implications for policymakers, business associations, and companies. For policymakers and institutional actors, the platform provides a tool to promote IS practices at the territorial level, supporting the creation of local and regional symbiotic networks. For companies, particularly small and medium-sized enterprises, MatchIS reduces the complexity associated with identifying and implementing IS opportunities by centralising information and providing access to specialised support services.

At the same time, the study highlights some critical aspects that may influence the effectiveness of digital platforms for IS. User engagement emerges as a key factor, as the success of the platform depends on the active participation of firms and the continuous updating of information. In addition, issues related to trust and data sharing remain relevant, despite the design choices aimed at mitigating them. The decision to externalise the

transactional phase may reduce these concerns, but it also implies that the platform does not directly control or monitor the outcomes of symbiotic exchanges.

Finally, it is important to acknowledge that the present study is limited by the absence of empirical validation. While the platform design is grounded in both literature and stakeholder insights, its actual effectiveness in facilitating IS remains to be tested. Future research should therefore focus on evaluating the performance of MatchIS, analysing user behaviour, the number and quality of established exchanges, and the resulting environmental and economic benefits.

6 Conclusions

The study has presented the design of MatchIS. Building on the literature review of IS barriers and the gaps of existing platforms, the study aimed at addressing these obstacles for the design of a new, open, and public IS platform. The platform explicitly incorporates different symbiotic roles, including producers, utilisers, and transformers, and offers tools to assist firms in evaluating the economic, legal, and environmental feasibility of potential exchanges. In conclusion, the MatchIS platform represents a promising step towards more effective digital support for IS. Its integrated design addresses several of the key barriers identified in the literature, while also opening new avenues for research on the role of digital technologies in enabling circular economy transitions. Given the recent launch of the platform, this study focuses on its design and conceptual development. The empirical evaluation of its effectiveness is left for future research, which will assess its impact on the adoption of IS practices and its ability to overcome the identified barriers.

7 Ethics Declaration

No ethical clearance for the research was required for this study.

8 AI Declaration

DeeplAI was used for language editing during the final review. Figure 1 was created with NapkinAI.

References

Akhtar, N., Bokhari, S.A., Martin, M.A., Saqib, Z., Khan, M.I., Mahmud, A., Zaman-ul-Haq, M. and Amir, S. (2022) “Uncovering Barriers for Industrial Symbiosis: Assessing Prospects for Eco-Industrialization through Small and Medium-Sized Enterprises in Developing Regions”, Sustainability, Vol. 14. No. 11

Antikainen, M., Uusitalo, T. and Kivikytö-Reponen, P. (2018) “Digitalisation as an Enabler of Circular Economy”, Procedia CIRP, Vol. 73, pp 45-49.

Benedict, M., Kosmol, L. and Esswein, W. (2018) “Designing Industrial Symbiosis Platforms: From Platform Ecosystems to Industrial Ecosystems”, The 22nd Pacific Asia Conference on Information Systems, June 28-29, Yokohama, Japan.

Bocken, N.M.P., de Pauw, I., Bakker, C. and van der Grinten, B. (2016) “Product Design and Business Model Strategies for a Circular Economy”, Journal of Industrial and Production Engineering, Vol. 33, No. 5, pp 308-320.

Chertow, M. and Ehrenfeld, J. (2012) “Organizing Self-Organizing Systems”, Journal of Industrial Ecology, Vol. 16, No. 1, pp 13-27.

Desrochers, P., (2004) “Industrial Symbiosis: the case for market coordination”, Journal of Cleaner Production, Vol. 12, No. 8-10, pp 1099-1110.

Domenech, T., Bleischwitz, R., Doranova, A., Panayotopoulos, D. and Roman, L. (2019) “Mapping Industrial Symbiosis Development in Europe_ typologies of networks, characteristics, performance and contribution to the Circular Economy”, Resources, Conservation and Recycling, Vol. 141, pp 76-98.

European Commission (2020), “Circular Economy Action Plan – for a Cleaner and More Competitive Europe”, Publications Office of the European Union.

Fraccascia, L., and Yazan, D. M. (2018) “The role of online information-sharing platforms on the performance of industrial symbiosis networks”, Resources, Conservation and Recycling, Vol. 136, pp 473-485.

Ghisellini, P., Cialani, C. and Ulgiati, S. (2016) “A Review on Circular Economy: The Expected Transition to a Balanced Interplay of Environmental and Economic Systems”, Journal of Cleaner Production, Vol. 114, pp 11-32.

Golev, A., Corder, G.D. and Giurco D.P. (2014) “Barriers to Industrial Symbiosis: Insights from the Use of a Maturity Grid”, Journal of Industrial Ecology, Vol. 19, No. 1, pp 141-153.

Herczeg, G., Akkerman, R. and Hauschild, M.Z. (2018) “Supply chain collaboration in industrial symbiosis networks”, Journal of Cleaner Production, Vol. 171, pp 1058-1067.

Krom, P., Piscicelli, L. and Frenken, K. (2022) “Digital Platforms for Industrial Symbiosis”, Journal of Innovation Economics & Management, Vol. 3, No. 39, pp 215-240.

Lippi, G., Evola, R.S. and Vesce, E. (2025) “Industrial symbiosis patterns in eco-industrial parks: A sectoral analysis based on global empirical data”, Journal of Cleaner Production, Vol. 529, 146807.

Lombardi, R. (2017) “Non-technical barriers to (and drivers for) the circular economy through industrial symbiosis: a practical input”, Economics and Policy of Energy and the Environment, Vol. 1, No. 2, pp 171-189.

Lybæk, R., Christensen, T B. and Thomsen, T.P. (2021) “Enhancing policies for deployment of Industrial symbiosis – What are the obstacles, drivers and future way forward?”, Journal of Cleaner Production, Vol. 280, 124351.

Maqbool, A.S., Mendez-Alva, F. and van Eetvelde, G. (2018) “An Assessment of European Information Technology Tools to Support Industrial Symbiosis”, Sustainability, Vol. 11, No. 1, 131.

Martin, M. and Eklund, M. (2011) “Improving the environmental performance of biofuels with industrial symbiosis”, Biomass Bioenergy, Vol. 35, No. 5, pp 1747-1755.

Mendez-Alva, F., Cervo, H., Krese, G. and van Eetvelde, G. (2021) “Industrial symbiosis profiles in energy-intensive industries: Sectoral insights from open databases”, Journal of Cleaner Production, Vol. 314, 128031.

Neves, A., Godina, R., G. Azevedo, S., and Matias, J. (2019) “Current Status, Emerging Challenges, and Future Prospects of Industrial Symbiosis in Portugal”, Sustainability, Vol. 11, No. 19, 5497.

Neves, A., Godina, R., G. Azevedo, S., and Matias, J. (2020) “A comprehensive review of industrial symbiosis”, Journal of Cleaner Production, Vol. 247, 119113

Oliveira Pavan, M.C., Soares Ramos, D., Yones Soares, M. and Carvalho, M.M. (2021) “Circular business models for bioelectricity: a value perspective for sugar-energy sector in Brazil”, Journal of Cleaner Production, Vol. 311, 127615.

Patricio, J., Kalmykova, Y.,Rosado, L., Cohen, L., Westin, A. and Gil, J. (2022) “Method for identifying industrial symbiosis opportunities”, Resources, Conservation and Recycling, Vol. 185, 106437.

Prosman, E.J. and Wæhrens B.V. (2019) “Managing waste quality in industrial symbiosis: insights on how to organize supplier integration”, Journal of Cleaner Production, Vol. 234, pp 113-123.

van Capelleveen, G., Amrit, C., and Yazan, D.M. (2018) “A Literature Survey of Information Systems Facilitating the Identification of Industrial Symbiosis”, Progress in IS, in: Otjacques B., Hitzelberger, P., Naumann S. and Wohlgemuth V. (ed.), From Science to Society, pages 155-169, Springer.

Velenturf, A.P.M. and Jensen, P.D. (2015) “Promoting Industrial Symbiosis: Using the Concept of Proximity to Explore Social Network Development”, Journal of Industrial Ecology, Vol. 20, No. 4, pp 700-709.

Wadström, C., Johansson, M. and Wallén, M. (2021) “A framework for studying outcomes in industrial symbiosis”, Renewable and Sustainable Energy Reviews, Vol. 151, 111526.

The Roles of Arts in Knowledge Creation in Urban Planning

Mervi Hokkanen
LUT University, Lahti, Finland

Abstract

A changing society challenges municipalities and their partners. Factors such as climate change, diversification of the regional structure and an ageing population require new approaches to community and urban planning. This research explores how values of arts and arts-based initiatives (ABIs), frameworks developed by Giovanni Schiuma, can contribute to knowledge creation in urban planning, a phase in which environmental and social challenges can be addressed and more sustainable urban environments and futures can be envisioned. Traditional management and planning principles focus merely on efficiency and technical knowledge. This does not recognise the complex reality of planning, where different understandings of valid knowledge coexist and compete. Diverse ways of knowing are required. Through arts, emotional knowledge can be integrated into work practices, enabling environmental challenges to be addressed from more multi-perspectival viewpoints. The study was conducted as a qualitative case study in a company specialising in urban planning. ABIs (such as visual storytelling methods which combine drawings, text and told stories) were used to co-create the research materials. In collaboration with the company's planners, the study examined their perceptions of their own work processes and explored the potential roles of ABIs in organisational everyday work practices. As preliminary results of the study, four operational models for everyday work practices were developed, through which the values of arts and arts-based initiatives (ABIs) support the achievement of the company's strategic objectives. Three of the proposed operational models relate to the collaborative ideation process and the fourth model is based on a reflexive framework. These models were aligned with four values of arts supporting personal inspiration, learning, and the strengthening of professional networks. The study provided practical insights for knowledge management by highlighting the importance of everyday work practices as a means of integrating and mobilising the organisation's collective knowledge.

Keywords: value of arts, arts-based initiatives (ABIs), urban planning, knowledge creation, knowledge management

Paper type: Academic Research Paper

1 Introduction

The Association of Finnish Cities and Municipalities has stated that a changing society brings challenges to municipalities and their partners. Factors such as climate change, diversification of the regional structure and an ageing population require new approaches to community and urban planning. Municipalities are required to engage in multidimensional change management and to collaborate with a wide range of stakeholders. (Jarva and Toivonen, 2020).

The built environment is produced through the cooperation of actors from several different fields and under the influence of societal processes (Hytönen, 2020). Finnish urban planning and land-use planning are hierarchical, top-down-guided processes that involve several levels of planning. Planning is a participatory process, involving not only public authorities and various experts but also those affected by the plan, such as residents, associations, and other stakeholders.

In the organisational structures of Finnish municipalities, a comprehensive-rationalist planning culture is still strongly present. This assumes that the state of society can be managed through comprehensive data collection methods and knowledge grounded in facts and objectivity. (Bäcklund and Mäntysalo, 2009). Such an interpretation of the nature of knowledge, worldview, and legitimate decision making (Bäcklund and Mäntysalo, 2009) does not recognise the complex reality of planning where different understandings of the determinants of good democracy, legitimate planning and the roles of different actors in producing and managing valid knowledge coexist and compete (Bäcklund and Mäntysalo, 2010). In a complex reality urban planners encounter conflicts between economics, environmental, and social interest that cannot be resolved through visionary imagery (Campbell, 1996). If the constraints set by the environment are taken seriously, the nature of planning changes (Hytönen, 2020). In planning, a new understanding of the conceptions of knowledge, worldview and legitimacy in decision-making is needed (Bäcklund and Mäntysalo, 2010).

In urban planning, the traditional management and planning principles focused merely on efficiency and technical knowledge are still present. However, this is not sufficient for solving complex problems or creating new value; instead, diverse ways of knowing are also required. Through the arts, emotional knowledge can be

integrated into work practices, enabling environmental challenges to be addressed from more multi-perspectival viewpoints. (Carlucci and Schiuma, 2018).

Arts-based initiatives (ABIs) have a unique character in the increase of knowledge and skill (Carlucci and Schiuma, 2018), and long-term impacts are achieved when they are integrated into everyday life (Taylor and Ladkin, 2009; Pässilä, Oikarinen and Harmaakorpi, 2015). In the knowledge processes underlying ABIs the senses and emotions are emphasized alongside logical-rational knowledge (Darsø, 2004; Taylor and Ladkin, 2009; Schiuma, 2011). Researchers study these themes under the fields of organisational aesthetics (Darsø, 2004; Taylor and Ladkin, 2009; Schiuma, 2011), sensuous knowing (Taylor and Hansen, 2005) and emotive knowing (Schiuma, 2011). Central and common to these approaches is the role of the senses in knowledge acquisition, as well as the significance of emotions and aesthetic experience in the interpretation of information and the creation of knowledge (Schiuma, 2011; Taylor and Ladkin, 2009).

Schiuma (2011) situates his approach to the roles of arts in knowledge within a body of scholarship that conceptualises human cognitive and emotive knowledge processes as deeply intertwined and mutually influential. As a result, human action and behaviour are grounded in both rational and emotive modes of cognition. Emotional knowledge processes activate individuals' aesthetic, sensory and experiential dimensions of knowing. Emotions and aesthetics are reciprocally related: aesthetic experiences generate emotion, while emotions shape aesthetic experiences and perceptions. ABIs are connected to emotional and energetic processes in individuals. Art has the ability to influence people's emotions and energy through the senses, and to evoke aesthetic experiences. Aesthetic experiences can shape the attitudes through which we interpret situations and decide how to act. The actions determine the quality and condition of the organisation's performance. (Schiuma, 2011). Through this operational logic, it is possible to influence not only to the rational dimension of the thinking but also various forms of business performance. Schiuma emphasises that integrating emotional knowledge processes into organisational practices is essential for postmodern management in an increasingly complex world.

Schiuma (2011) has been exploring the benefits of arts through the framework of the Arts Value Matrix he developed. The Arts Value Matrix structures the types of value that can be achieved by integrating arts-based initiatives into individual and organizational learning and development. The Arts Value Matrix examines individual and organizational development at low, medium and high levels of intensity. The Arts Value Matrix covers nine categories, which are entertainment, galvanising, inspiration, reputation, environment, learning and development, investment, networking, and transformation.

To make these connections between arts and business visible, Schiuma (2011) draws on a knowledge-based understanding of organisational operations, where all the organisation's resources, activities, and products can be interpreted and classified as being the result of knowledge and expertise. Knowledge resources are constituted by people and their relationships, as well as by structures that can be either tangible - such as material artifacts, physical environments, and infrastructure – or intangible, such as organisational culture and identity. What distinguishes Schiuma's conception of knowledge resources from other knowledge management literature is that he also recognizes tangible resources as part of knowledge resources. Tangible resources contribute directly to organisational capabilities, while also influencing intangible resources through the networks of interdependencies among organisational resources. Both tangible and intangible resources are linked to the organisational and structural properties of the organisation. (Schiuma, 2011).

Schiuma also draws on performance management literature in developing the Arts Value Map framework, which elucidates how the operational logic of the arts can be translated into business performance outcomes. This framework enables the visualisation of relevant cause-and-effect relationships. ABIs can function as a platform for learning or as a means of developing an organisation's knowledge resources. These, in turn, influence the organisation's operational and dynamic capabilities, which affect organisational processes and, consequently, the achievement of the values the organisation aspires to create. ABIs influence people and structures in a distinctive, context-specific, temporal, and multilevel way. (Schiuma, 2011).

2 Research Objective

Interest in bringing artists and arts into planning processes has increased (Wiberg, 2022), but this is still an area of research that is just emerging (Carpenter and Horvath, 2022). This research explores how ABIs can contribute to knowledge creation in urban planning, a phase in which environmental and social challenges can be addressed and more sustainable urban environments and futures can be envisioned. The objective of this study is to investigate how the ABIs may be integrated into the everyday work practices of a Finnish company operating in the field of urban planning, while simultaneously supporting the attainment of the company's

strategic objectives. In addition, the study aims to develop operational models for embedding the values of arts and ABIs into everyday work practices.

Accordingly, the research question is formulated as follows:

How can ABIs be embedded with staff and managers in everyday organisational practices in ways that activate arts-driven value-creation mechanisms and thereby support the implementation of the company's strategic priorities?

3 Data, Method and Analysis

This study employs a qualitative research approach and is designed as a case study (Yin, 2018) in which ABIs were used within the everyday context of knowledge-creation practices. ABIs (such as visual storytelling methods which combine drawings, text and told stories) were used to co-create the research materials. In collaboration with the company's planners, the study examined their perceptions of their own work processes and explored the potential roles of ABIs in organisational everyday work practices.

3.1 Data collection process

Research data were collected in 2022 in a company operating in the urban planning sector. The data collection process consisted of two workshops, an orientation meeting prior the second workshops, and a concluding discussion at the end of process. The data were co-produced with the company's staff (12 people) and managers (4 people), all who work on planning projects. In the first workshop, managers described the strategic objectives of the company and the everyday work practices related to these objectives. These objectives and practices were identified and structured using the Arts Value Map developed by Schiuma (2011). Prior to the second workshop, an orientation meeting was held for staff and managers. The purpose of this meeting was to develop a preliminary shared understanding and to outline the themes to be addressed in the subsequent workshop. In the second workshop, staff and managers jointly described everyday work practices that managers had identified as strategically important. These practices were then analysed in terms of what functions well and what presents challenges. Following this description and analysis, participants engaged in an ideation phase focusing on how the Arts Values developed by Schiuma could be integrated into everyday work practices. The data collection process concluded with a discussion with managers, during which they reflected on the events and insights generated throughout the process.

3.2 Data and methods

In the study, various visual storytelling methods were used and combined as data collecting methods. Maps represent one of the oldest forms of non-verbal communication and serve as a powerful tool for visualising reality, relationships between entities, and the construction of meaning (Carlucci, Schiuma and Linzalone, 2017). For this reason, the workshop series began with the creation of an Arts Value Map. This framework is based on mapping practices and helps to understand how ABIs can be translated into business performance outcomes. (Carlucci et al., 2017).

Drawing functions as a catalyst for collective meaning-making as well as for sharing emotions and experiences (Ward and Shortt, 2018). In the data collection process, various drawing-based methods were employed. Participants drew visual stories depicting their everyday work practices. To support storytelling, guiding questions derived from the Work Stories storytelling method (Pässilä et al., 2013) were used. As well as the visualisation and narration of their own Work Stories, participants also listened to others' stories and created visual notes based on them, applying an evocative visual storytelling approach. Through this method, information is collected, the process is documented and tacit knowledge is given expression. The method also enables the visualisation of relationships between entities and supports the creation of shared understanding. (Benmergui, 2019). In addition, Work Stories were documented using the (maga)zine method. This is a flexible and three-dimensional approach that supports the visualisation of complex issues. (Biagioli, Pässilä and Owens, 2021).

In data collection methods, visual material was supported and completed through verbal means, including writing and interviews. In this way, the limitations of visual material were addressed with verbal explanations, while aspects that are difficult to express in words were more easily communicated through visualisation.

The research material consists of drawings, written texts, and verbal descriptions and discussions. This material captures both the present and desired future activities of the company. The material was analysed using Polkinghorne's (1995) narrative analysis approach. The collected data were classified and categorised, and these categories were made visible through the narratives constructed from the described Work Stories. Based

on the analysis of these narratives, proposals for operational models were developed for experimentation in the company. The analytical categories leading to the operational models were derived from the objectives, the knowledge resources, and the value of arts identified in the first workshop, as well as from the descriptions and analysis of current-state knowledge and the values of arts discussed in the second workshop. These categories were grounded in the frameworks of the Arts Value Map, the Arts Value Matrix and the concept of knowledge resources developed by Schiuma (2011).

4 Results

This study is work-in-progress. As preliminary findings of the study, four proposed operational models were developed for experimentation within everyday work practices. These models were aligned with four values of arts supporting personal inspiration, learning, and the strengthening of professional networks.

The strategic objectives and values identified in the study emphasised the company's aspiration to become a dream workplace for its employees as well as a forerunner in the field of urban planning. From the perspective of being a dream workplace, collaborative ideation was identified as a key everyday work practice. Ideation is a routine task in a planning office and a central component of the planner's professional role. The managers viewed this practice as particularly important, as it enables employees to experience success, a sense of community and commitment to the company. To achieve the strategic objective of forerunnership, informal and spontaneous encounters – such as coffee breaks and casual interactions in the office – were considered important contexts for sharing experiences and perspectives related to work. The following section presents three proposed operational models for integrating the values of arts into collaborative ideation, and one proposed operational model for integrating the values of arts into informal and spontaneous encounters.

Three arts-related values were linked to collaborative ideation to develop the forms of knowledge required in this practice. The first value was the galvanising value of the arts, which materialises when art can engage individuals in an intense, liberating, and meaningful experience that generates enthusiasm, joy, and energy (Schiuma, 2011). Schiuma (2011) further explains that the galvanising value of the arts can enhance people's motivation and commitment, as well as knowledge creation and sharing, creativity, and innovativeness. Through the galvanising value of the arts, the intention was to support the development of employees' personal competences and skills, such as imagination and creativity.

The first proposed operational model integrates the galvanising value of the arts into collaborative ideation as an everyday work practice. In this model, arts-based methods engaging multiple senses are incorporated into the ideation phase of collaborative design processes. During the workshops, merely ideating the use of different arts-based methods within the co-design process generated enthusiasm among participants, suggesting that the actual implementation of these methods is likely to produce a galvanising effect as well. These methods include the use of sound recording, photography, drawing, and the production of documentary films. Employees and managers can adapt, vary, and further develop these methods on a project-by-project basis to support ideation.

A second value of arts to be integrated into collaborative ideation was the inspirational value of the arts. According to Schiuma (2011), this value involves offering new perspectives, stimulating ideas, and inspiring action. The inspirational value of the arts is realised when it guides individuals toward self-reflection and deeper understanding, thereby transforming thinking and behaviour (Schiuma, 2011). Through the inspirational value of the arts, the intention was to support employees' personal competences and skills, as well as to develop the organisation's collaborative ideation knowledge management process, which forms part of the organisation's structural resources.

The second proposed operational model (Figure 1) integrates the inspirational value of the arts into collaborative ideation as an everyday work practice. This model is based on the creative process described by Taylor (2011). The inspirational value of the arts is realised within the creative process outlined by Taylor (2011), which emphasises processual progression and the development of deeper understanding as a foundation for idea generation. In addition to fostering individual deeper understanding, the process establishes a stepwise work practice that structures collaborative ideation.

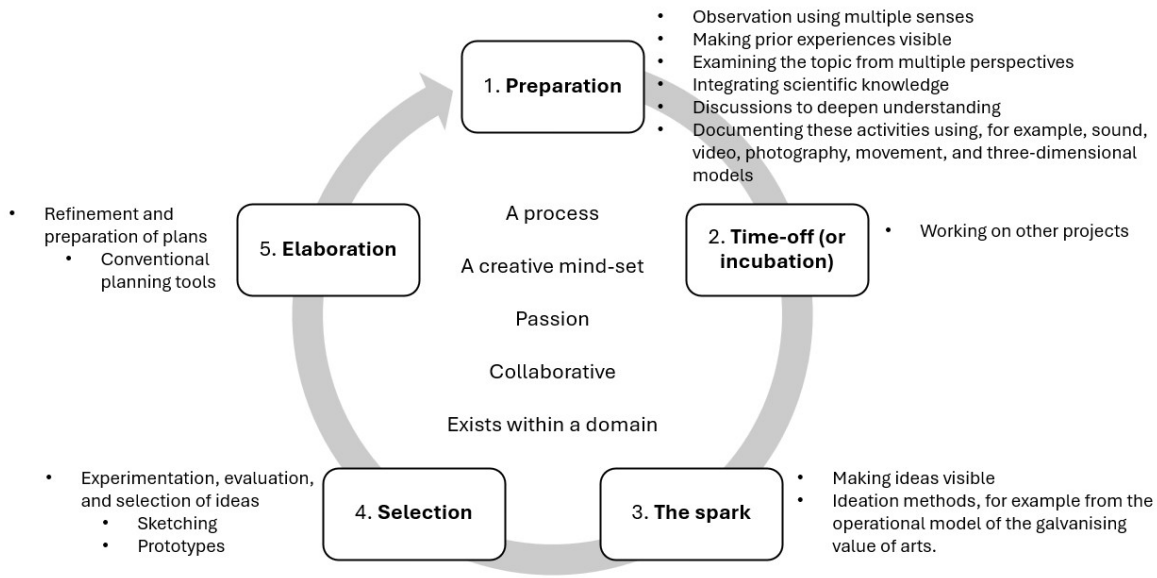


Figure 1: The proposed operational model: integrating the arts inspiration value into collaborative ideation.

The third value of arts to be integrated into collaborative ideation was the value of arts as networking. The value of the arts in networking is connected to the formation of shared understanding through dialogue, as well as to the activation and support of collaboration. Within this value domain, the arts can generate collective effects by bringing people together and stimulating dialogue that helps to build relationships as a foundation for cooperation (Schiuma, 2011). In the employees' narratives, challenges related to the quality of relational knowledge within multidisciplinary planning teams were highlighted. The ideation process defined and described by the managers was explicitly collaborative ideation, in which individual knowledge is transformed into shared knowledge.

The third proposed operational model (Figure 2) integrates the value of arts of networking into collaborative ideation as an everyday work practice. This model draws on the study by Sorsa et al. (2018) on the transformation of individual experience into collective collaboration and knowledge, which reflects the characteristics of the arts networking value described by Schiuma (2011). According to Sorsa et al. (2018), the foundation for collective dialogue lies in awareness of oneself and of others. They further argue that a materially mediated aesthetic experience is essential when bridging individual experience and collective collaboration (Sorsa et al., 2018).

The process examined by Sorsa et al. (2018) consists broadly of the following phases: self-awareness; doing individually; listening and taking the perspective of others; and doing together. In this process, aesthetic experience and material artifacts function as modes of communication, particularly in situations where issues are difficult to articulate verbally (Sorsa et al., 2018). The aim of the proposed operational model is to increase self-awareness and, through this, promote mutual understanding, attentive listening, the expression of diverse perspectives and approaches, improved collaboration, and the creation of shared understanding among participants regarding the project. These outcomes are intended to foster mutual trust and shared understanding, thereby improving the effectiveness of collaborative ideation. The aesthetic experience and materials employed may consist of previously ideated multisensory or other task-appropriate arts-based methods.

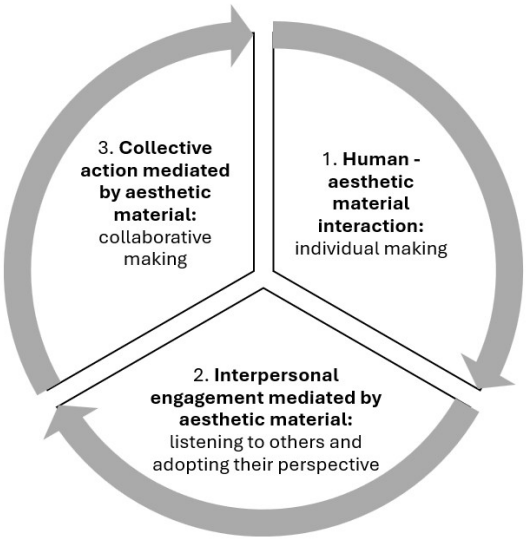


Figure 2: The proposed operational model: integrating the arts networking value into collaborative ideation.

The fourth proposed operational model and values of arts are linked to the strategic objective of forerunnership, for which informal and spontaneous everyday conversations were identified as important contexts for achievement. Four values of arts are associated with this objective: galvanising, inspiration, learning and development, and networking. The value of arts of learning and development activates and sustains experiential learning mechanisms in which the arts function through reflective learning processes. By practicing artistic skills or engaging with artistic competences, individuals' capabilities can be enhanced (Schiuma, 2011). The value of arts of networking is connected to the formation of shared understanding through dialogue, as well as to the activation and support of collaboration (Schiuma, 2011). The other values of arts have been presented in earlier sections.

Through these values of arts, the aim is to develop multiple organisational knowledge resources. The managers identified facilities, office culture, and trust as key forms of knowledge. In their development ideas, the employees emphasised the importance of personal experiences, interaction and collaboration, spaces and tools, and leadership. The significant and development-oriented forms of knowledge resources were found to relate to all categories of knowledge assets defined by Schiuma (2011): individual, relational, tangible, and intangible structures.

The fourth proposed operational model focuses on the more systematic integration of informal and spontaneous knowledge exchange into everyday work practices. This model draws on Biagioli's (2017) descriptions of the nature and significance of informal situations in organisations. Informal situations play an important role in organisational knowledge formation, particularly in relation to innovation. They are also significant for cultural understanding and organisational change. Informal situations provide contexts in which the transfer of tacit knowledge is especially activated, as formal practices do not guide action in these settings. In such situations, tacit knowledge is transferred and transformed through gestures, playfulness, informal learning, and embodied feelings associated with individual beliefs. (Biagioli, 2017).

The proposed operational model (Figure 3) is built around the reflexive model developed by Pässilä (2012) and aims to make visible and to construct the understanding that the case company perceives as emerging in spontaneous and informal situations. Through the values of arts, everyday experiences are rendered visible, multiple perspectives related to these experiences are shared, situations are interpreted in a polyphonic manner, new ways of acting are imagined, and novel practices are tested in action. Reflexive work practice should be given its own place within the company's structures and processes of knowledge creation and cultural understanding.

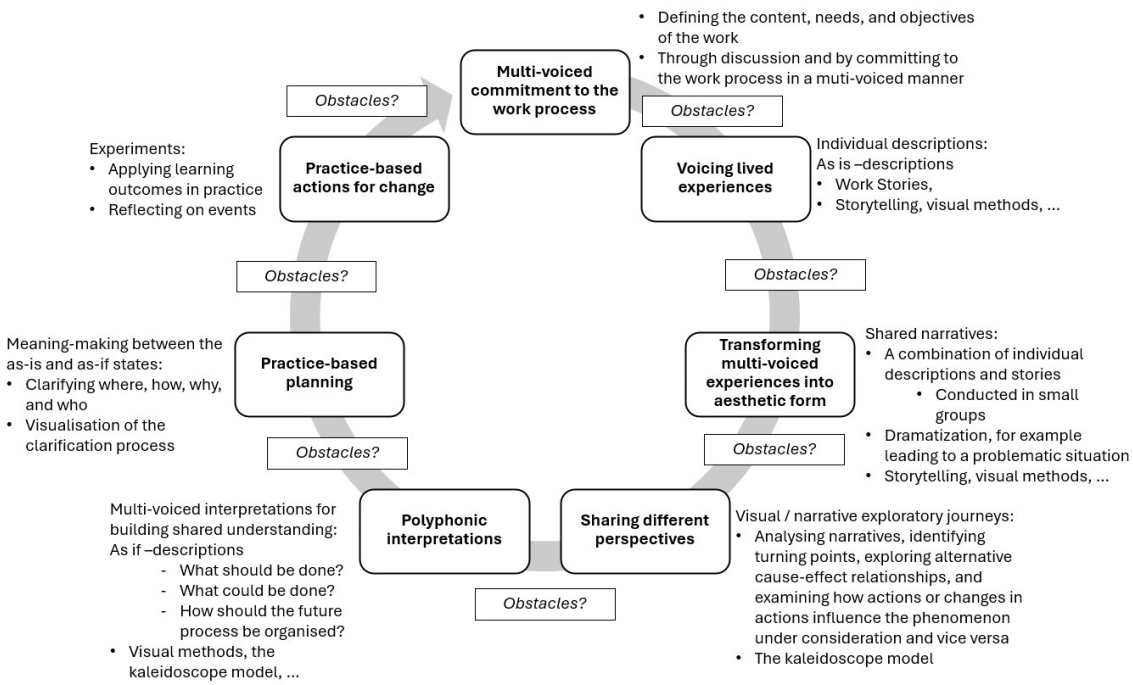


Figure 3: The proposed operational model: integrating the arts galvanising, inspiration, learning and development, and networking values into reflexive model.

The study provided practical insights for knowledge management by highlighting the importance of everyday work practices as a means of integrating and mobilising the organisation's collective knowledge. The frameworks developed by Schiuma (2011) support the structuring of how the values of arts and ABIs can be embedded within the organisation's strategic activities. To align values of arts and the ABIs that promote them with strategically significant everyday work practices, a shared description and analysis of these practices involving the entire organisation—including both staff and managers—is required.

Work practices comprise of multiple phases in which different forms of knowledge embedded in individuals and organisational structures are combined. The findings of this study indicate that ABIs are most effectively designed as operational models that integrate multiple values of arts, thereby enabling the simultaneous support and development of diverse knowledge forms. From a managerial perspective, it is also noteworthy that, based on the findings, informal situations within the organisation emerged as important arenas for knowledge creation.

5 Conclusions

This study examines the roles that the arts can play in urban planning, particularly from the perspective of knowledge creation. As preliminary results of the study, four operational models for everyday work practices were developed, through which the values of arts and arts-based initiatives (ABIs) support the achievement of the company's strategic objectives. Three of the proposed operational models relate to the collaborative ideation process, in which the galvanising, inspiration, and networking values of the arts are used to support employees' imagination and creativity, deeper self-reflection and understanding, and the transformation of individual experiences into collective collaboration and knowledge. The objective of these models is to achieve the strategic goal of becoming a dream workplace by supporting the development of employees' knowledge and competencies.

The fourth proposed operational model is based on a reflexive framework, in which the galvanising, inspiration, learning and development, and networking values of the arts are employed to render tacit knowledge visible and to generate new understanding. Through this model, knowledge creation occurring in informal situations within the company is provided with a dedicated platform, thereby supporting the attainment of the strategic objective of forerunnership.

The study is situated in Finland; however, planning structures and methods share similarities with those of many other European countries (Rantanen and Rajaniemi, 2020). This case study explored the roles of arts from

the perspective of an urban planning company and examined how arts influences professionals' everyday work practices.

In the next phase of the research, the focus will shift to investigating the roles that arts may play in articulating vulnerable residents' relationships with the city and in bringing their perspectives into the planning process. Differentiation in participation is evident both within residential areas and in participation in urban planning more broadly. A case study from Helsinki indicates that current planning practices have limited success in reaching socially vulnerable or otherwise marginalised groups. (Sjöblom and Niitamo, 2020). Future research will also find it relevant to examine the roles of third-sector organisations and municipal urban organisations, as well as how knowledge produced through arts-based methods can be utilised.

In addition to making emotional knowledge visible (Schiuma, 2011), arts (as an aesthetic approach) also have the potential to shape our aesthetic perceptions and our relationship with the environment (Lehtinen, 2021).

6 Ethics declaration

The ethical clearance for the research referred to in this paper was not required.

7 AI declaration

AI-based tool assistance of Enterprise Microsoft Copilot 365 was utilised in translating text.

References

Benmergui, R. (2019) Making the Invisible visible: Evocative Visual Storytelling & Visceral Liminal Documentation of Process & Sense-Making, *Beyond Text – Art-Based Method for Research*, *Assessment and Evaluation*, eds. Benmergui, R., Owens, A. and Pässilä, A., [online] *Beyond text july2019*

Biagioli, M. (2017) The Case for Informal Spaces in the Workplace, *Integrating Art and Creativity into Business Practice*, eds. Schiuma, G. and Lerro, A. IGI Global, Her shey, USA.

Biagioli, M., Pässilä, A. and Owens, A. (2021) The zine method as a form of qualitative analysis, *Beyond Text – Learning through Arts-Based Research*, eds. Adams, J. and Owens, A., Intellect Ltd, Bristol, UK.

Bäcklund, P. and Mäntysalo, R. (2009) Yhdyskuntasuunnittelun teorioiden kehitys ja asukkaiden osallistumisen paikka, *Terra*, Vol. 121, No. 1, pp 1-18.

Bäcklund, P. and Mäntysalo, R. (2010) Agonism and institutional ambiguity: Ideas on democracy and the role of participation in the development of planning theory and practice - the case of Finland, *Planning Theory*, Vol. 9, No. 4, pp 333-350.

Carlucci, D. and Schiuma, G. (2018) The power of the arts in business, *Journal of Business Research*, Vol. 85, April, pp 342-347.

Carlucci, D., Schiuma, G. ja Linzalone, R. (2017) A Fuzzu Cognitive Map-Based Approach for Understanding Value Creation Mechanisms of Artistic Interventions, *Integrating Art and Creativity into Business Practice*, eds. Schiuma G. and Lerro A. IGI Global.

Carpenter, J and Horvath, C. (2022) Co-Creation and the City: Arts-Based Methods and Participatory Approaches in Urban Planning. *Urban Planning*, Vol. 7, No. 3, pp 311-314.

Darsø, L. (2004) *Artful Creation: Learning-tales of Arts-in-business*, Narayana Press, Gylling, Denmark.

Jarva, A. and Toivonen, L. (2020) Kumppanuuden monet muodot. Esimerkkejä yhteistyöhön perustuvista kaavoituskäytännöistä, Kuntaliitto.

Hytönen, J. (2020) Hyvän suunnittelun muuttuvat periaatteet, *Yhdyskuntasuunnittelu*, Vol. 58, No. 2, pp 87-90.

Lehtinen, S. (2021) Aesthetic Sustainability, *Situating Sustainability: A Handbook of Contexts and Concepts*, eds. Toivanen, R. and Krieg, P. Helsinki: Helsinki University Press, pp 255–268.

Polkinghorne, D. E. (1995) Narrative configuration in qualitative analysis, *Life history and narrative*, eds. Hatch J. A. and Wisniewski, R. Falmer, London.

Pässilä, A. (2012) Reflexive model of research-based theatre. Processing innovation at the crossroads of theatre, reflection and practice-based innovation activities. (Doctoral dissertation, Lappeenranta University of Technology). Acta Universitatis Lappeenrantaensis 492. [online] Reflexive model of research-based theatre – processing innovation at the crossroads of theatre, reflection and practice-based innovation activities

Pässilä, A. H., Oikarinen, T., and Harmaakorpi, V. (2015) Collective voicing as a reflexive practice, *Management Learning*, Vol. 46, No. 1, February, pp 67–86.

Pässilä, A., Oikarinen, T. and Kallio, A. (2013) Creating dialogue by storytelling, *Journal of Workplace Learning*, Vol. 25, No. 3, pp 159-177.

Rantanen, A. and Rajaniemi, J. (2020) Urban planning in the post-zoning era: From Hierarchy to self-organisation in the reform of the Finnish Land Use and Building Act, *Environment and Planning B: Urban Analytics and City Science*, Vol. 47, No. 2, February 2020, pp 321-335.

Schiuma, G. (2011) *The Value of Arts for Business*, Cambridge University Press, New York.

Sjöblom, J. and Niitamo, A. (2020). The intermediating role of municipal urban planners in online discussions with citizens, *Planning Theory & Practice*, Vol. 21, No. 5, pp 710–726.

Sorsa, V., Merkkiniemi, H., Endrissat, N. and Islam, G. (2018) Little less conversation, little more action: Musical intervention as aesthetic material communication, *Journal of Business Research*, Vol. 85, April, pp. 365-374.

Taylor, S. S. (2011) *Leadership Craft*, *Leadership Art*, Palgrave Macmillan, New York.

Taylor, S. S. & Hansen, H. (2005) Finding Form: Looking at the Field of Organizational Aesthetics, *Journal of Management Studies*, Vol. 42, No. 6, pp 1211-1231.

Taylor, S. S. and Ladkin, D. (2009) Understanding Arts-Based Methods in Managerial Development, *Academy of Management Learning & Education*, Vol. 8, No. 1, pp 55-69.

Ward, J. and Shortt, H. (2018) Drawing, *The Sage Handbook of Qualitative Business and Management Research Methods: Methods and Challenges*, eds. Cassell, C., Cun liffe, A. L. & Grandy, G. Sage Publications Ltd.

Wiberg, S. (2022) Planning With Art: Artistic Involvement Initiated by Public Authorities in Sweden, *Urban Planning*, Vol. 7, No. 3, pp 394-404.

Yin, K. R. (2018) *Case Study Research and Applications: Design and Methods*, Sixth Edition, Sage Publications, Inc, Los Angeles.

Using Intelligent Knowledge to Offer Digital Services: a Lean Startup Interpretation

Clara Filosa, Lara Agostini, Anna Nosella
Department of Management and Engineering, University of Padova, Vicenza, Italy

Abstract

Digital servitization (DS) enables manufacturing firms to transition from product-centric offerings to digitally enabled product-service systems (PSSs). Despite its strategic potential, DS represents a challenging transformation for incumbent manufacturers, characterized by technological, data-driven, and business model uncertainty. While prior research conceptualizes DS as a strategic and organizational shift, limited attention has been paid to the process mechanisms through which firms generate and leverage knowledge to navigate this transformation. In particular, the role of Intelligent Knowledge (IK) and the mechanisms through which firms act remain insufficiently understood.

This study investigates how IK evolves throughout the DS journey and how the Lean Startup approach helps firms manage uncertainty during this process. To do so, the research draws on a longitudinal multiple-case study of four global B2B manufacturing firms undergoing DS transformations.

Our findings reveal a three-stage pathway through which firms progressively develop and mobilize IK. In the first stage, experimentation enables firms to identify relevant customer problems and achieve problem-solution fit. In the second stage, accumulated insights support refinement and the achievement of PSS-market fit. In the third stage, IK expands to ecosystem-level knowledge, enabling outcome-based services and problem-solution fit.

The study contributes to the DS literature by introducing a knowledge-based perspective on the transformation process and by showing how Lean Startup experimentation functions as an adaptive mechanism to navigate the journey. For managers, the findings provide a structured roadmap for DS, highlighting how staged experimentation and progressive IK development can guide the transition toward scalable PSSs.

Keywords: Digital Servitization, Intelligent Knowledge, Lean Startup approach, Multiple case study.

Paper type: Academic Research Paper

1 Introduction

Digital servitization (DS), the transition from selling standalone physical products to offering digitally enabled and integrated product-service systems (PSSs), is increasingly recognized as a key pathway to generate competitive advantage. By leveraging digital technologies, firms can deliver adaptive solutions that continuously respond to evolving customer needs (Frank et al., 2019; Kohtamäki et al., 2020; Stonig et al., 2022). Despite these opportunities, DS represents a major challenge for incumbent manufacturers: organizational inertia, legacy infrastructures, and product-centric cultures often hinder the transition (Baines et al., 2020; Favoretto et al., 2022; Huikkola et al., 2022).

Beyond technological adoption, DS requires firms to generate, interpret, and act upon new forms of knowledge derived from connected products, digital platforms, and customer interactions. However, despite the central role of data and knowledge in DS, limited attention has been devoted to understanding how this evolves and supports decision-making throughout the servitization journey. In particular, little is known about how incumbents generate and use intelligent knowledge (IK), defined as a form of knowledge that emerges from processing and interpreting data and supports decision-making (North et al., 2018). As a result, scholars have called for new perspectives capable of capturing the dynamic and emergent nature of DS as it unfolds in practice (Kohtamäki et al., 2022).

In this context, the Lean Startup approach offers a promising lens for understanding how firms navigate uncertainty while progressively generating and acting upon new knowledge. Originally developed to guide startups operating under conditions of high uncertainty, Lean Startup emphasizes hypothesis-driven experimentation, iterative learning, and continuous customer feedback (Ries, 2011). Recent research suggests that Lean Startup may help incumbents manage innovation under uncertainty (Blank & Eckhardt, 2024; Ghezzi & Cavallo, 2020). However, how Lean Startup operates within DS transformations remains underexplored.

To address these gaps, this paper investigates how IK and Lean Startup evolve during DS transformations. Drawing on a longitudinal qualitative study of four B2B manufacturers undergoing DS transformations, we address two interconnected research questions:

RQ1: How does Intelligent Knowledge evolve along the digital servitization journey?

RQ2: How can the Lean Startup approach help firms navigate uncertainty throughout this journey?

Our findings reveal a three-stage pathway through which IK progressively evolves from early product-level insights to broader ecosystem-oriented knowledge. Across these stages, Lean Startup provides an adaptive approach that enables firms to translate emerging knowledge into iterative service development and BM transformation under uncertainty.

2 Theoretical Background

2.1 Digital servitization as a knowledge-intensive transformation

DS refers to the transformation through which manufacturing firms integrate digital technologies into their PSS offerings to develop data-enabled services and new value propositions (Frank et al., 2019; Kohtamäki et al., 2020).

At its core, DS requires profound business model (BM) transformation: manufacturers must reconfigure value creation, delivery, and capture mechanisms to support the development and provision of digitally enabled PSSs (Kohtamäki et al., 2020; Paiola et al., 2024).

This transformation is inherently knowledge intensive. Connected products and digital platforms enable firms to gather continuous data across the product lifecycle, supporting predictive services, remote monitoring, and data-driven optimization of customer operations (Baines et al., 2020). As a result, successful DS initiatives depend on new knowledge management capabilities, including data analytics, digital infrastructures, and cross-functional collaboration (Huikkola et al., 2022; Paiola & Gebauer, 2020).

Importantly, DS also expands knowledge processes beyond firm boundaries. Digital services often require collaboration with customers and partners, leading to data-driven service ecosystems in which knowledge is co-created and shared among multiple actors (Sjödin et al., 2020). Despite growing scholarly attention, however, limited research has examined how knowledge is generated and mobilized as DS transformations unfold, highlighting the need for perspectives capable of capturing their iterative and emergent nature (Sjödin et al., 2020).

2.2 Intelligent knowledge enabled by digital technologies

In DS contexts, digital technologies do not merely support operations but function as infrastructures for knowledge generation. IoT sensors and analytics platforms continuously generate data on product usage and performance. When processed and interpreted, these data generate insights that inform service delivery. Following North et al. (2018), we conceptualize this digitally enabled form of knowledge that supports data-driven decision-making as Intelligent Knowledge (IK).

In DS, IK is continuous and actionable. IK flows in real time, providing manufacturers with an uninterrupted stream of insights into product performance and customer behaviour (Solem et al., 2022). Moreover, IK is not merely descriptive but prescriptive, enabling firms to trigger proactive service interventions (i.e., predictive maintenance alerts, remote performance optimizations, automated anomaly responses) that translate insights into value (Kohtamäki et al., 2022; Sjödin et al., 2020).

IK enables DS by revealing actual rather than assumed customer usage patterns, by continuously identifying opportunities for value-adding interventions, and by enabling remote monitoring, automated diagnostics, and evidence-based service scheduling (Kohtamäki et al., 2022; Tian et al., 2022).

However, IK also introduces significant uncertainty and challenges. The huge volume, variety, and velocity of data generated by connected industrial assets create analytical and interpretive burdens for organizations (Hanelt et al., 2021). Raw data are not insights: transforming operational signals into actionable service knowledge requires significant investment in analytical capabilities, organizational routines, and governance structures for knowledge mobilization across functions (Kohtamäki et al., 2022). Moreover, as digital technologies evolve rapidly and become increasingly integrated, firms remain uncertain about the most appropriate digital infrastructure and technology combinations.

Consequently, IK simultaneously enables and complicates DS, producing a foundational tension that demands structured organizational responses.

2.3 Lean startup approach to manage digital servitization transformation

DS exposes incumbent manufacturers to multiple and interrelated forms of uncertainty. Technological uncertainty stems from the rapid evolution of digital technologies such as IoT, cloud infrastructures, and AI (Frank et al., 2019; Kohtamäki et al., 2020). Firms also face data-driven uncertainty regarding which technologies

to adopt and how to interpret the growing volume of data (Sjödin et al., 2020). Finally, DS introduces BM uncertainty, as manufacturers must experiment with new ways of creating, delivering, and capturing value through digitally enabled PSSs (Baines et al., 2020; Paiola & Gebauer, 2020).

In such contexts, the Lean Startup approach provides a useful framework for managing innovation under uncertainty. Originally developed for early-stage ventures, it emphasizes hypothesis testing, experimentation, and validated learning to progressively refine value propositions and BMs (Ries, 2011; Ghezzi, 2019). Recent studies suggest that these mechanisms can also support incumbent firms undergoing complex innovation-driven transformations (Ghezzi & Cavallo, 2020; Blank & Eckhardt, 2024).

A key mechanism within this approach is the Think-Make-Check cycle (Alvarez, 2017), through which firms formulate hypotheses about customer problems (Think), develop simplified experimental solutions (Make), and evaluate market feedback to validate or revise their assumptions (Check).

Despite these conceptual synergies, how Lean Startup approach operates within DS transformations remains empirically underexplored, particularly regarding how experimentation interacts with the generation and mobilization of knowledge along the servitization journey.

3 Methodology

To address the RQs, we adopted a qualitative multiple case study approach suitable for exploring complex phenomena (Yin, 2009). Multiple case studies enable cross-case comparison and robust analysis (Siggelkow, 2007). The investigation spans multiple time points (2023-2025), allowing the identification of changes and dynamics both retrospectively and in real time (Pettigrew, 1990).

Four global B2B manufacturers were selected based on the following criteria: being incumbent manufacturers actively engaged in DS projects, operating at different scales, and willing to provide regular access to data and participate in interviews and meetings. The sample includes Alpha (large firm, energy infrastructure), Beta (large firm, professional ovens), Gamma (medium firm, industrial thermal solutions), and Delta (medium firm, packaging machines), operating in diverse industrial contexts while sharing a common commitment to DS transformation.

Primary data were collected through semi-structured interviews and meetings with informants actively engaged in DS projects (Pettigrew, 1990). Secondary sources (i.e., company websites, internal presentations, and official documents) were incorporated to triangulate primary findings and ensure narrative completeness (Yin, 2009).

Interviews were recorded, transcribed, and rigorously coded to identify first-order categories grounded in informants' language. Connections among first-order categories were explored iteratively to derive second-order themes corresponding to distinct theoretical notions. Finally, a coding scheme aligned empirical findings with Lean Startup and IK.

4 Findings

Our cross-case analysis reveals that the DS process unfolds through three iterative stages, namely Hypothesis Framing and Initial Experiments, Solution Validation, and Extended and Ecosystem-Oriented Experiments, each characterized by a distinct configuration of IK and Lean Startup mechanisms. Across stages, IK progressively evolves from product-level operational data to ecosystem-level strategic intelligence, while Lean Startup mechanisms shift from basic hypothesis testing to complex partner co-development. Together, these dynamics illuminate how incumbent manufacturers adapt and transform their BMs under the compounding uncertainties of DS, successfully navigating the journey.

4.1 Stage 1: Hypothesis framing and initial experiments

In Stage 1, manufacturers initiate their DS journey by formulating hypotheses about customer problems and developing early-stage digital service configurations to test them on a limited scale. The emphasis is on learning what works and demonstrating value to selected customers. Firms begin the journey, without yet committing to full-scale transformation.

4.1.1 Intelligent knowledge evolution

In Stage 1, IK is product-centric and internally oriented. Firms invest in IoT sensors to capture operational data from their physical assets, generating a foundational layer of knowledge about product performance and usage conditions.

At Beta, connectivity is introduced to *"detect parameters such as temperature settings and water consumption to understand customer habits, optimizing and improving the product itself"* (Chief Information Officer - CIO). The data are initially accessible only to internal technicians, serving as a learning instrument about actual product usage rather than a vehicle for service delivery. Similarly, Delta equip its machines with sensors since 2017, noting that *"critical components were monitored to collect data throughout the asset lifecycle"* (Chief Operations Officer - COO).

Beyond product-level data, IK in Stage 1 begins to incorporate nascent customer knowledge, gathered through structured but preliminary interactions. Gamma leverage IoT data collected since 2015 to hypothesize digital service proposals: *"We actively involved selected customers through a targeted survey to validate our hypotheses on digital services and gather valuable feedback before moving into development"* (Service Specialist). Delta present an early-stage service concept at a major international trade fair, using customer reactions as a first-order signal of market interest: *"In 2018, at the most important trade fair for our sector, we presented the early-stage hypothesized concept to our key customers, testing their interest and gathering essential insights for its subsequent development"* (COO). These interactions represent the first steps in converting product-level IK into customer-relevant service intelligence, establishing the informational foundation for PSSs development.

4.1.2 Lean startup mechanisms

The dominant Lean Startup mechanism in Stage 1 is the Think-Make-Check cycle. In the Think phase, firms articulate and structure hypotheses about customer problems before committing to development. Beta hypothesize that lack of visibility on product use hampered quality improvements and service innovation. Gamma identifies that customers fail to recognize the value of previously free services and begins designing service contract: *"The idea was to make customers recognize the value and make them pay for the supporting activities"* (Service Specialist). Alpha hypothesize that customers need long-term operational reliability rather than transactional support: *"The time had come to expand the offer and think about how to constantly guarantee the continuity of the operational activities of our assets"* (Executive). This problem framing directs subsequent experiments.

In the Make phase, firms develop early-stage digital service configurations integrating physical products, connectivity, and service support. Alpha develops a data platform providing AI- and IoT-enabled asset diagnostics: *"IoT inspection data, photos and videos are collected in a component inspection repository; there is an assessment of the asset status, and recommendations are elaborated using AI and analytics tools"* (Executive). In parallel, Alpha develops a B2B e-commerce platform for spare parts and an advanced analytics suite for predictive maintenance and power forecasting. Beta extends its web and mobile monitoring platform - previously used only by internal technicians - to customers, offering remote oven monitoring and fleet management as a complimentary service: *"We combined the digitalized product with a service: a web and a mobile platforms that allow us to control the ovens at any time, even remotely, in real time"* (Chief Experience Officer - CXO). Gamma integrates a remote monitoring platform into its boiler service contracts *"to provide analysis at three levels: for the end customer, for service partners, and for the company which has full control and access to all information"* (CEO). Delta introduces preventive maintenance and anomaly detection services: *"We combined the tele-assistance and monitoring of the machine data... it makes sense to combine remote assistance and data monitoring"* (Operations Manager).

In the Check phase, PSSs are tested. To this end, all solutions are deliberately limited in scope. Beta offers its platform for free initially, treating it as a learning instrument rather than a revenue vehicle. Alpha positions its digital services as value-added components of existing contracts. Gamma and Delta conduct pilot projects with selected clients. As Gamma's President acknowledges: *"We began with prototype phases to thoroughly understand the needs and determine how to proceed"*. Delta's COO confirmes the same logic: *"The digital services developed underwent a testing phase with technicians on selected clients in Italy and Europe before expanding the offering"*. Thanks to this approach, firms gather feedback: customer responses to the new value propositions provide empirical signals about which service configurations generate value.

4.2 Stage 2: Solution validation

Following initial experiments, firms enter a solution validation phase in which they refine digital infrastructures and service functionalities. The emphasis shifts from testing whether PSSs create value to ensuring they do so effectively and in a scalable way.

4.2.1 Intelligent knowledge evolution

In Stage 2, IK evolves from product-level data toward richer forms of intelligence, including customer-behavioural and operational insights, which inform service refinement and progressively support BM optimization.

Beta illustrates this transition: its platform evolves from basic monitoring to an AI-enabled system that analyses usage patterns and automates operations: "*The platforms use AI to analyse usage patterns, provide consumption notifications, and automate certain operations*" (CXO). Similarly, Alpha consolidates fragmented digital platforms into a single interface to improve usability based on customer interaction data: "*We moved from a multichannel system to a single-sign-on platform that customers found easier to access*" (Executive). Gamma's GECO platform (GEnerators Characterization and Optimization) represents a further step in this integration process. The platform combines operational data, predictive analytics, and service interfaces within a single digital environment that "*supports boiler selection and predictive maintenance*" (Service Manager). By consolidating previously fragmented service and operational data into a unified analytical infrastructure, GECO enables multi-level insights for customers, service partners, and the focal firm simultaneously. Delta integrates edge computing capabilities with existing cloud and AI systems: "*While AI was initially employed to analyse data collected from IoT sensors in the cloud, it is now applied directly at the machine level - on the edge. This allows for faster, localized data processing and real-time decision-making, reducing latency and enabling immediate responses to anomalies*" (COO). This shift moves knowledge generation closer to operations, enabling faster responses to anomalies.

4.2.1 Lean startup mechanisms

The dominant Lean Startup mechanism in Stage 2 is validated learning applied to PSSs refinement. Customer feedback accumulated at the end of Stage 1 becomes the empirical basis for systematic service and infrastructure optimization.

Alpha unifies its fragmented platforms into a single cohesive interface with single sign-on access, directly responding to customer-reported usability barriers. Beta merges its separate web and mobile applications into one synchronized platform: "*Using the same code, we produced an application with two interfaces, one for the mobile version and one for the online one. In contrast to the past, when they were handled individually and might display different data, we now satisfy customer preference to have integrated applications that are always aligned*" (CIO). Beyond infrastructure consolidation, Beta introduces a new generation of smart ovens equipped with IoT sensors, AI-driven anomaly detection, and a connected control panel enabling direct remote intervention, exemplifying how product innovation and service innovation become tightly interwoven through validated learning. Gamma consolidates previously separated service functionalities into the GECO platform, making all insights and performance reports accessible through a unified, user-friendly interface that "*enhances transparency, usability, and customer engagement*" (Service Manager). Delta advances its AI capabilities to the edge level, streamlining service delivery and incorporating functionalities "*reflecting insights gathered from customer feedback*" (COO).

Next to value offering optimization, Stage 2 marks the critical transition from complimentary or experimental service access to sustainable recurring revenue models. Beta provides the clearest illustration: having offered its platform for free in Stage 1, it begins charging for the enhanced, feature-rich service, communicating to customers that "*to maintain it and ensure functionality, there is a cost the company must bear in terms of people and resources dedicated to the project... taking advantage of this type of feature implies a cost; to continue accessing it, a subscription is required*" (CXO). This transition is legitimized by the demonstrated value: customers who experienced the platform usefulness are willing to pay for its continued, enhanced availability. Gamma pushes multi-year service packages: "*The company is now pushing to sell packages with a duration of 3-5 years. In case of subscription for more than one year, the annual or monthly price remains fixed*" (Service Specialist). Delta moves toward subscription-based contracts coupled with machine sales: "*We are moving towards subscription-based contracts. The sales of the machines are coupled with a yearly-basis subscription fee*" (Operations Manager).

The validation implemented in Stage 2 provides the foundation for the more ambitious ecosystem-level experimentation that characterizes Stage 3.

4.3 Stage 3: Extended and ecosystem-oriented experiments

In Stage 3, manufacturers leverage the consolidated capabilities and validated BM configurations developed in Stages 1 and 2 to expand their service portfolios and engage in broader ecosystem collaborations, where value creation, delivery, and capture are no longer confined within the firm's boundaries but distributed across a

network of interconnected stakeholders. Empirical evidence for this stage mainly derives from Alpha and Beta, which progressed more rapidly in the DS journey; the same trajectory is nonetheless expected for medium-sized firms, albeit at a slower pace.

4.3.1 Intelligent knowledge evolution

In Stage 3, IK expands to ecosystem-level knowledge: it not only informs operational decisions within firms, but also shapes how value is co-created across partners and customers. As digital platforms open to external actors, the knowledge generated through connected assets becomes a shared resource that supports coordinated service delivery. For instance, Alpha progressively opens its digital platform to external partners and begins to monetize operational data, recognizing that *"We are part of an ecosystem where we need to connect our data to others and, the other way around, connect others into our ecosystem"* (Executive). Therefore, IK evolves from an internal analytical resource into a mechanism that supports ecosystem coordination and contractual governance: the continuous knowledge generated by connected assets provides the objective basis for measuring value delivery and managing long-term service relationships.

4.3.2 Lean startup mechanisms

Stage 3 reactivates the Think-Make-Check cycle at the ecosystem level, applying Lean Startup logic to a new form of experiments where the hypotheses being tested now require engagement with external stakeholders.

In the Think phase, firms engage in hypothesis-driven ideation about ecosystem-level PSSs. Alpha hypothesizes that opening its digital platform to third-party partners could unlock new service opportunities and revenue streams: *"At the industry level, several players operate around the company. We recognized the importance of becoming an active part of this broader ecosystem, staying open, avoiding silos, and fostering coordinated integration among stakeholders"* (Executive). This hypothesis involves framing data access and system interoperability as foundational to future value propositions. Beta begins questioning the limitations of the product sales model and hypothesizes an alternative BM based on usage: *"Some key customers indicated that their interest in purchasing ovens was starting to decline, while there was growing interest in rental programs"* (CXO). Critically, Beta's Think phase extends beyond customers to include dealers: *"We began moving in this direction by involving not only customers, to better understand how to design and develop this offering, but also the dealers who act as intermediaries between us and the customers. This step was crucial: engaging the dealers was essential to maintain their trust and ensure their collaboration, consequently safeguarding our market position"* (CXO).

In the Make phase, firms translate ecosystem hypotheses into new PSSs involving complex usage- and outcome-based contracts. Alpha operationalizes its ecosystem strategy by exposing APIs and enabling integration with customers' internal systems and third-party providers, co-developing services with ecosystem actors and embedding its solutions into broader customer workflows. In parallel, Alpha introduces performance-based contracts: *"We offer what we call 'power solutions': not just the uptime, but we also guarantee the output by collaborating with partners at the industry level"* (Executive). Beta operationalizes its rental hypothesis by launching a five-year oven rental model with pay-per-use pricing logic. The offering is built on a dedicated digital platform through which customers independently configure their rental, specifying expected usage that directly determine pricing: *"Through a dedicated platform, the customer independently selects the oven, including all desired features and optional accessories. Customers specify expected daily usage hours, which determine pricing. The system then automatically identifies the authorized resellers in the customer's area who have joined the program and generates a summary displaying the monthly installation fee, the required monthly payment, and the total value of services provided"* (Project Manager). The subscription fee includes not only the equipment itself but also predictive maintenance and full technical support throughout the contract period.

In the Check phase, firms engage in ecosystem-level test: rather than selected customers as in Stage 1, Stage 3 piloting involves assembling and testing multi-actor service delivery configurations, assessing not only customer acceptance but also partner coordination effectiveness. Beta's rental program is explicitly structured as a controlled ecosystem pilot: *"The oven rental project started as a pilot... we have currently selected a few dealers who are collaborating on the project, and now the rental service is available in specific regions. It is not yet a project that has been expanded nationwide"* (CXO). Alpha applies the same graduated approach to ecosystem engagement, explicitly recognizing the need for incremental partner involvement: *"We need to proceed through steps, involving the partners gradually and understanding the respective benefits"* (Executive).

Throughout the DS journey, firms have evolved from single-stream product sales to complex, data-enabled, ecosystem-distributed value capture systems, a transformation made possible through the iterative application of Lean Startup experimentation guided by progressively sophisticated IK.

5 Discussion and Conclusions

This study investigated how incumbent manufacturers navigate the DS process, introducing IK as a key enabler and presenting the Lean Startup approach as an interpretative lens. By analysing four B2B manufacturers, we identify a three-stage pathway in which IK and Lean Startup mechanisms co-evolve.

In Step 1, firms explore DS opportunities by testing hypotheses about customer problems. At this stage, experiments primarily aim to achieve problem-solution fit: firms identify relevant customer problems and test whether proposed digital solutions can effectively address them (Blank & Eckhardt, 2024), leveraging early forms of IK based on product usage data and initial customer interactions.

In Step 2, firms consolidate insights emerging from early experiments and refine PSSs accordingly. The accumulation and analysis of customer interaction data enable the development of more advanced forms of IK that guides PSSs improvements in terms of both digital infrastructure and service functions. Hence, firms move toward PSS-market fit, achieved since integrated PSSs demonstrate sustained adoption and willingness to pay.

In Step 3, firms build on the knowledge accumulated in earlier stages to engage ecosystem partners. Experiments continue but focus on expanding BMs within a wider ecosystem.

Overall, our findings show that IK is a key element in the DS journey, which can be interpreted through a Lean startup lens. Experiments generate knowledge about customer problems and service value, enabling the transition from problem-solution fit to PSS-market fit.

From a theoretical perspective, the study shows that Lean Startup functions as an adaptive learning mechanism for incumbents undergoing DS (Blank & Eckhardt, 2024; Ghezzi & Cavallo, 2020), thus expanding the applicability of this approach beyond startup contexts. Furthermore, it contributes to DS research by explaining how experiments and knowledge generation drive digital PSSs development (Kohtamäki et al., 2020; Sjödin et al., 2020).

From a managerial standpoint, our study shows that experiments enable firms to navigate DS uncertainty: first validating problem-solution fit through experimentation, then refining services to achieve PSS-market fit, and finally expanding digital offerings through ecosystem collaboration.

This study has limitations, including the small sample. Future research could examine broader contexts and factors influencing DS experimentation.

6 Ethics Declaration

Ethical approval was not required for this study.

7 AI Declaration

No AI tools were used in the preparation of this manuscript.

References

Alvarez, C. (2017). Lean customer development: Building products your customers will buy. O'Reilly Media, Inc.

Baines, T., Bigdeli, A. Z., Sousa, R., & Schroeder, A. (2020). Framing the servitization transformation process: A model to understand and facilitate the servitization journey. International Journal of Production Economics, 221, 107463.

Blank, S., & Eckhardt, J. T. (2024). The Lean Startup as an Actionable Theory of Entrepreneurship. Journal of Management, 50(8), 3012–3034. https://doi.org/10.1177/01492063231168095

Eisenhardt, K. M., & Graebner, M. E. (2007). Theory Building From Cases: Opportunities And Challenges. Academy of Management Journal, 50(1), 25–32. https://doi.org/10.5465/amj.2007.24160888

Favoretto, C., Mendes, G. H., Oliveira, M. G., Cauchick-Miguel, P. A., & Coreynen, W. (2022). From servitization to digital servitization: How digitalization transforms companies' transition towards services. Industrial Marketing Management, 102, 104–121.

Frank, A. G., Dalenogare, L. S., & Ayala, N. F. (2019). Industry 4.0 technologies: Implementation patterns in manufacturing companies. International Journal of Production Economics, 210, 15–26.

Ghezzi, A. (2019). Digital startups and the adoption and implementation of Lean Startup Approaches: Effectuation, Bricolage and Opportunity Creation in practice. Technological Forecasting and Social Change, 146, 945–960.

Ghezzi, A., & Cavallo, A. (2020). Agile business model innovation in digital entrepreneurship: Lean startup approaches. Journal of Business Research, 110, 519–537.

Hanelt, A., Bohnsack, R., Marz, D., & Antunes Marante, C. (2021). A Systematic Review of the Literature on Digital Transformation: Insights and Implications for Strategy and Organizational Change. Journal of Management Studies, 58(5), 1159–1197. https://doi.org/10.1111/joms.12639

Huikkola, T., Kohtamäki, M., & Ylimäki, J. (2022). Becoming a smart solution provider: Reconfiguring a product manufacturer's strategic capabilities and processes to facilitate business model innovation. Technovation, 118, 102498.

Kohtamäki, M., Parida, V., Patel, P. C., & Gebauer, H. (2020). The relationship between digitalization and servitization: The role of servitization in capturing the financial potential of digitalization. Technological Forecasting and Social Change, 151, 119804.

Kohtamäki, M., Rabetino, R., Parida, V., Sjödin, D., & Henneberg, S. (2022). Managing digital servitization toward smart solutions: Framing the connections between technologies, business models, and ecosystems. In Industrial Marketing Management (Vol. 105, pp. 253–267). Elsevier.

Langley, D. J., van Doorn, J., Ng, I. C., Stieglitz, S., Lazovik, A., & Boonstra, A. (2021). The Internet of Everything: Smart things and their impact on business models. Journal of Business Research, 122, 853–863.

North, K., Maier, R., & Haas, O. (2018). Knowledge management in digital change. New findings and practical cases.

Paiola, M., & Gebauer, H. (2020). Internet of things technologies, digital servitization and business model innovation in BtoB manufacturing firms. Industrial Marketing Management, 89, 245–264.

Paiola, M., Grandinetti, R., Kowalkowski, C., & Rapaccini, M. (2024). Digital servitization strategies and business model innovation: The role of knowledge-intensive business services. Journal of Engineering and Technology Management, 74, 101846.

Pettigrew, A. M. (1990). Longitudinal Field Research on Change: Theory and Practice. Organization Science, 1(3), 267–292. https://doi.org/10.1287/orsc.1.3.267

Ries, E. (2011). The lean startup: How today's entrepreneurs use continuous innovation to create radically successful businesses. Crown Currency.

Siggelkow, N. (2007). Persuasion With Case Studies. Academy of Management Journal, 50(1), 20–24. https://doi.org/10.5465/amj.2007.24160882

Sjödin, D., Parida, V., Jovanovic, M., & Visnjic, I. (2020). Value creation and value capture alignment in business model innovation: A process view on outcome-based business models. Journal of Product Innovation Management, 37(2), 158–183.

Solem, B. A. A., Kohtamäki, M., Parida, V., & Brekke, T. (2022). Untangling service design routines for digital servitization: Empirical insights of smart PSS in maritime industry. Journal of Manufacturing Technology Management, 33(4), 717–740.

Stonig, J., Schmid, T., & Müller-Stewens, G. (2022). From product system to ecosystem: How firms adapt to provide an integrated value proposition. Strategic Management Journal, 43(9), 1927–1957. https://doi.org/10.1002/smj.3390

Tian, J., Coreynen, W., Matthyssens, P., & Shen, L. (2022). Platform-based servitization and business model adaptation by established manufacturers. Technovation, 118, 102222.

Yin, R. K. (2009). Case study research: Design and methods (Vol. 5). sage.

Leading Intelligent Knowledge Management for Data-Driven Sustainable Innovation: Micro-Practices that Build Human–AI Complementarities

Vittoria Magrelli[1], Isotta Vazzoler[1], Dario N. Palmucci[2], Luca Giraldi[2]
[1]University Gabriele D'Annunzio, Chieti-Pescara, Italy
[2]University of Turin, Turin, Italy

Abstract

Intelligent Knowledge Management (IKM) represents a shift from static knowledge repositories toward dynamic socio-technical arrangements in which data, human expertise, and algorithmic inference are continuously combined to support organisational action. While existing research has examined knowledge creation processes, analytics capabilities, and the implications of AI for decision-making, less attention has been paid to how leaders make these elements work together in practice—particularly in the context of sustainability-oriented innovation. Addressing this gap, this study explores how leadership is enacted through everyday practices that transform AI- and data-driven outputs into actionable organisational knowledge. Adopting a practice-theory perspective, the study conceptualises leadership as a set of situated micro-practices through which problems are framed, insights are interpreted, and learning is stabilised over time. Empirically, it draws on a qualitative, comparative case study of organisations engaged in AI-enabled knowledge work for sustainability. The research combines semi-structured interviews with senior leaders and analytics professionals, analysis of internal artefacts such as knowledge governance protocols and AI guidelines, and observational material where accessible. An abductive analytical approach is used to iteratively connect empirical patterns with theoretical insights from knowledge management, organisational learning, and human–AI augmentation literatures. The findings are expected to show that IKM becomes effective not through technology adoption alone, but through leadership practices that makes knowledge usable. These include framing sustainability-relevant problems in ways that guide analytical attention, curating and translating AI-generated outputs for diverse organisational audiences, fostering productive contestation that balances critical scrutiny with integration of experiential knowledge, and embedding insights into routines that sustain learning beyond individual projects. From a recombinant innovation perspective, these practices enable the continuous recombination of heterogeneous knowledge—spanning past and present, human and machine—thereby supporting cumulative rather than episodic innovation. By foregrounding leadership as a governance mechanism of intelligent knowledge, the study contributes to bridging knowledge management, digital transformation, and sustainability research, offering both theoretical refinement and practical guidance for organisations operating in complex, data-rich environments.

Keywords: Intelligence Knowledge Management; Leadership, Data-driven innovation; Sustainable innovation; Organizational learning; Recombinant innovation.

Paper type: Academic Research Paper

1 Introduction

Intelligent Knowledge Management (IKM) can be understood as a shift from knowledge management systems conceived as repositories toward socio-technical arrangements that continuously transform data, experience, and interpretation into knowledge that can be used (Alavi & Leidner, 2001; Nonaka, 1994). This shift is not primarily technological. It is managerial: it concerns how organisations frame problems, mobilise attention, coordinate across functions, and stabilise learning routines while integrating AI, advanced analytics, and natural language technologies into everyday work (Davenport & Kirby, 2016; Raisch & Krakowski, 2021). In this sense, IKM is best approached as a capability for orchestrating human judgement and machine-supported inference, rather than as the adoption of an additional tool (Raisch & Krakowski, 2021). The concept draws on and extends classical knowledge management theory, which distinguished between explicit and tacit knowledge and emphasised the social processes through which knowledge is created, shared, and institutionalised within organisations (Nonaka & Takeuchi, 1995; Grant, 1996). What distinguishes IKM from earlier frameworks is the integration of machine learning, natural language processing, and large-scale data analytics into these social processes, creating what some scholars describe as hybrid epistemic arrangements in which human and algorithmic inference are mutually constitutive (Faraj et al., 2018). This integration introduces both new opportunities and new governance challenges. On the opportunity side, AI-enabled systems can surface patterns in large, heterogeneous datasets that would be invisible to unaided human attention, extending the reach of organisational sensemaking and enabling more systematic exploration of innovation possibilities (Berente et al.,

2021). On the governance side, the very abundance of data and the opacity of algorithmic models can generate interpretive overload, introduce new forms of bias, and create a false sense of precision that undermines rather than supports deliberative decision-making (Cukier et al., 2021; Brynjolfsson & McAfee, 2017). Sustainability-oriented innovation increasingly depends on data-driven exploration and recombination: organisations search for new ways to reduce waste and resource intensity, redesign value propositions, and experiment with more durable forms of organising (Geissdoerfer et al., 2017; Bocken et al., 2014). The transition toward circular economy models, for instance, generates vast streams of product lifecycle, resource flow, and consumer behaviour data that require intelligent knowledge systems to be translated into actionable insights (Ellen MacArthur Foundation, 2013; Lacy & Rutqvist, 2015). Similarly, organisations pursuing science-based emissions targets or responsible supply chain commitments must develop the capacity to combine environmental performance data with operational and strategic knowledge in ways that inform innovation decisions across business units and value chain boundaries (Nambisan et al., 2017). Importantly, recombination is not only about producing novelty from "new" data; it also depends on how organisations retrieve, reinterpret, and redeploy dispersed knowledge over time. Recent process work on recombinant innovation shows how firms can mobilise "old" knowledge through distinct phases—recovering, reviving, and renewing—so that it becomes a platform for strategic renewal rather than a constraint (Magrelli et al., 2025). This reinforces a managerial point that is central to IKM: insights do not travel automatically, and AI outputs do not become organisational knowledge by themselves. They become actionable only when leaders shape the conditions under which they are interpreted, challenged, and translated into decisions and routines (Weick, 1995; Argote & Miron-Spektor, 2011). Therefore, as highlighted by recent studies, this makes leadership central (Palmucci & Ferraris, 2026; Palmucci et al., 2025): not as generic support for digital transformation, but as a set of micro-practices that govern how knowledge is produced, circulated, and used (Felin et al., 2015).

2 Research Gap and Purpose

Existing work offers strong foundations on (i) organisational knowledge creation and the socio-technical nature of knowledge systems (Nonaka, 1994; Alavi & Leidner, 2001), (ii) analytics capabilities and their organisational implications (Akter et al., 2016; Ferraris et al., 2019), and (iii) the automation–augmentation tension introduced by AI (Raisch & Krakowski, 2021). Scholarship on dynamic capabilities and knowledge-based views of the firm has further established that the ability to integrate, build, and reconfigure knowledge resources is a primary driver of competitive advantage and organisational resilience (Teece et al., 1997; Kogut & Zander, 1992). More recently, research on digital transformation has examined how AI and data analytics reshape organisational processes, competitive positioning, and value creation (Vial, 2019; Zuboff, 2019). What remains under-specified is how leaders *actually* make IKM work in practice when the strategic intent is sustainable innovation: how they build human–AI complementarities, prevent "data abundance" from turning into interpretive fragmentation, and convert analytics into cumulative organisational learning (Argote & Miron-Spektor, 2011; Weick, 1995). Building on a recombinant innovation lens, this also means explaining how leaders enable the organisation to recombine heterogeneous knowledge—new and old, human and machine-mediated—into a coherent innovation trajectory (Magrelli et al., 2025).

This study addresses the following research question: How do leaders enact Intelligent Knowledge Management in their day-to-day practice to enable data-driven sustainable innovation? This question is grounded in a practice-theory perspective that treats leadership not as a set of traits or styles but as something that is accomplished through situated, concrete activities—what strategy scholars have called "strategy-as-practice" and what organisation theorists describe as the microfoundations of collective action (Jarzabkowski, 2005; Felin et al., 2015). A practice lens is particularly suited to study IKM because the translation of data and AI outputs into organisational knowledge is fundamentally a social achievement: it happens in meetings, briefing sessions, review processes, and informal exchanges through which leaders frame what counts as relevant evidence, who is authorised to challenge a model's outputs, and how analytically derived insights are connected to organisational memory and strategy (Orlikowski, 2002). The empirical focus is on leadership micro-practices that (a) frame and prioritise sustainability-relevant innovation problems (Abdulmuhsin et al., 2025), (b) organise interpretation and contestation of AI- and analytics-based insights so that they become shared understandings rather than isolated outputs, (c) sustain cross-functional and cross-boundary knowledge flows that keep innovation efforts coordinated, and (d) institutionalise learning into routines and decision processes that endure beyond individual projects (Felin et al., 2015; Argote & Miron-Spektor, 2011).

3 Methodology

The study follows a qualitative, comparative case design (Eisenhardt, 1989). It investigates 10 organisations selected through purposive, theory-driven sampling to maximise variation in (i) firm size, (ii) sector, and (iii) geographical setting, while keeping cases comparable in terms of AI-enabled knowledge work for sustainability. The target organisations are European-headquartered firms (primarily Italy and other EU countries), where sustainability reporting and innovation programmes are increasingly institutionalised, and where comparable regulatory and market pressures support cross-case comparison. We focus mainly on medium-sized and large organisations (typically >250 employees)—as they are more likely to have formalised knowledge governance, dedicated analytics/AI roles, and routinised innovation processes—while including a limited number of smaller firms as contrast cases where such routines are still emerging. Sectorally, cases will be drawn from data-rich and/or resource-intensive industries in which sustainability-oriented innovation is salient (e.g., manufacturing and industrial supply chains, logistics/retail operations, energy/utilities, and consumer-facing circular economy domains such as agri-food/fashion). Eligible cases must have introduced AI- and analytics-enabled knowledge processes explicitly linked to sustainability-oriented innovation initiatives (e.g., circular experimentation, resource optimisation, sustainable supply-chain redesign, or data-enabled service innovation) (Bocken et al., 2014; Geissdoerfer et al., 2017).

Data collection combines semi-structured interviews with senior leaders and key knowledge/analytics roles, internal artefacts (knowledge protocols, AI use guidelines, analytics routines, training materials), and observational material where accessible. Interviews will follow a topic guide designed to elicit detailed accounts of how leaders respond to specific episodes in which AI- or analytics-generated insights were contested, reinterpreted, or integrated into innovation decisions. This approach draws on the critical incident technique (Flanagan, 1954) as a way of grounding leadership accounts in specific behavioural episodes rather than general attitudes or self-reports. Internal artefacts will be analysed as materialised traces of knowledge governance: they reveal how organisations formalise (or fail to formalise) the conditions under which data and AI outputs are to be interpreted and acted upon (Carlile, 2004). Observational material, where accessible, will allow direct observation of knowledge practices in situ, providing a triangulation point for interview and document data (Yin, 2018). The analysis will use an abductive approach to connect emergent leadership practices with mechanisms suggested by socio-technical KM theory, organisational learning, and human–AI augmentation perspectives (Alavi & Leidner, 2001; Raisch & Krakowski, 2021). Abductive analysis involves moving iteratively between empirical observations and theoretical concepts, allowing the study to both leverage existing frameworks and generate new theoretical insights where the data reveals patterns that existing theories do not adequately explain (Timmermans & Tavory, 2012). The coding process will combine deductive codes derived from the theoretical framework (e.g., framing, curation, contestation, institutionalisation) with inductive codes emerging from the data, following established protocols for qualitative case research (Gioia et al., 2013).

4 Expected Results and Contributions

The study aims to produce a practice-grounded model of IKM leadership micro-practices that explains how leaders make intelligent knowledge usable for sustainability-oriented innovation. Anticipated mechanisms include: (1) problem framing and attention design, through which leaders define which sustainability tensions deserve analytic investment and experimentation (Weick, 1995); (2) curation and translation, through which AI- and analytics-based outputs are rendered interpretable and relevant for heterogeneous stakeholders; (3) productive contestation, through which leaders create safe spaces to challenge model-driven insights and integrate them with experiential knowledge, avoiding both blind trust and defensive rejection (Raisch & Krakowski, 2021); and (4) institutionalising learning, through which insights are embedded into routines, cross-functional interfaces, and repeatable innovation practices (Argote & Miron-Spektor, 2011). A recombinant innovation perspective sharpens the contribution by showing how leaders enable recombination across time and levels—mobilising dispersed knowledge assets and connecting them to present data streams and AI-supported inferences—so that innovation becomes cumulative rather than episodic (Magrelli et al., 2025). The expected contribution is to reposition leadership as a governance mechanism of intelligent knowledge in data-driven sustainable innovation, showing how human judgement and AI can be designed as complements in everyday managerial work (Felin et al., 2015; Davenport & Kirby, 2016). Beyond the theoretical contribution, the study is expected to generate practical implications for managers and organisations navigating the integration of AI and analytics into innovation processes. Specifically, the resulting model of IKM micro-practices can inform the design of leadership development programmes, knowledge governance frameworks, and AI deployment guidelines that position human expertise and machine intelligence as complementary rather than competing sources of organisational capability (Brynjolfsson & McAfee, 2017; Berente et al., 2021). The study also

contributes to the IFKAD community's interest in intelligent knowledge as a resource for sustainable organising by providing an empirically grounded account of how data-driven knowledge practices can be institutionally designed to serve long-term sustainability goals rather than short-term efficiency gains (Calabrese et al., 2019). In doing so, it responds to calls for research that bridges knowledge management, innovation studies, and sustainability science in ways that are actionable for practitioners operating in complex, data-rich environments (Lichtenthaler, 2019).

5 Ethics Declaration

This paper reports an ongoing qualitative study based on semi-structured interviews, complemented by internal documents and observation where available. The research is subject to institutional ethical review and adheres to standard academic guidelines, ensuring informed consent, voluntary participation, confidentiality, and anonymisation of all data. Participants receive clear information prior to data collection, in line with European data protection standards. No sensitive personal data or vulnerable populations are involved, and no deceptive or high-risk procedures are used. Therefore, no additional ethical approval is required beyond standard institutional review.

6 AI Declaration

Artificial intelligence tools were used in a limited and non-systematic manner during the preparation of this paper. Their use was restricted to minor language refinement and English proofreading, without influencing the core ideas, analysis, or conclusions. All content, arguments, and interpretations remain the sole responsibility of the authors. AI assistance was applied only for fine-tuning purposes to improve clarity, readability, and overall linguistic accuracy.

References

Abdulmuhsin, A. A., Hussein, H. D., AL-Abrrow, H., Masa'deh, R. E., & Alkhwaldi, A. F. (2025). Impact of artificial intelligence and knowledge management on proactive green innovation: the moderating role of trust and sustainability. Asia-Pacific Journal of Business Administration, 17(3), 765-795.

Akter, S., Wamba, S. F., Gunasekaran, A., Dubey, R., & Childe, S. J. (2016). How to improve firm performance using big data analytics capability and business strategy alignment? *International Journal of Production Economics*, 182, 113–131.

Alavi, M., & Leidner, D. E. (2001). Review: Knowledge management and knowledge management systems: Conceptual foundations and research issues. *MIS Quarterly*, 25(1), 107–136.

Argote, L., & Miron-Spektor, E. (2011). Organizational learning: From experience to knowledge. *Organization Science*, 22(5), 1123–1137.

Bocken, N. M. P., Short, S. W., Rana, P., & Evans, S. (2014). A literature and practice review to develop sustainable business model archetypes. *Journal of Cleaner Production*, 65, 42–56.

Davenport, T. H., & Kirby, J. (2016). *Only humans need apply: Winners and losers in the age of smart machines*. New York: Harper Business.

Eisenhardt, K. M. (1989). Building theories from case study research. *Academy of Management Review*, 14(4), 532–550.

Felin, T., Foss, N. J., & Ployhart, R. E. (2015). The microfoundations movement in strategy and organization theory. *Academy of Management Annals*, 9(1), 575–632.

Ferraris, A., Mazzoleni, A., Devalle, A., & Couturier, J. (2019). Big data analytics capabilities and knowledge management: Impact on firm performance. *Management Decision*, 57(8), 1923–1936.

Geissdoerfer, M., Savaget, P., Bocken, N. M. P., & Hultink, E. J. (2017). The Circular Economy: A new sustainability paradigm? *Journal of Cleaner Production*, 143, 757–768.

Magrelli, V., Kotlar, J., De Massis, A., Frattini, F., & Petruzzelli, A. M. (2025). Making the Old New Again Through the Process of Recombinant Innovation. *Journal of Product Innovation Management*. In press

Nonaka, I. (1994). A dynamic theory of organizational knowledge creation. *Organization Science*, 5(1), 14–37.

Palmucci, D. N., Ferraris, A. (2026). "Leading innovation at Zucchetti. An action research to validate a new definition of STP leadership". *European Journal of Innovation Management*. 1-21.

Palmucci, D. N., Giovando, G., Vincurova, Z. (2025). "The post-covid era: digital leadership, organizational performance and employee motivation". *Management Decision*.

Raisch, S., & Krakowski, S. (2021). Artificial intelligence and management: The automation–augmentation paradox. *Academy of Management Review*, 46(1), 192–210.

Weick, K. E. (1995). *Sensemaking in organizations*. Thousand Oaks, CA: Sage.

Berente, N., Gu, B., Recker, J., & Santhanam, R. (2021). Managing artificial intelligence. *MIS Quarterly*, 45(3), 1433–1450.

Brynjolfsson, E., & McAfee, A. (2017). *Machine, platform, crowd: Harnessing our digital future*. New York: W. W. Norton & Company.

Calabrese, A., Costa, R., Levialdi, N., & Menichini, T. (2019). Integrating sustainability into strategic decision-making: A fuzzy AHP method for the selection of relevant sustainability issues. *Technological Forecasting and Social Change*, 139, 155–168.

Carlile, P. R. (2004). Transferring, translating, and transforming: An integrative framework for managing knowledge across boundaries. *Organization Science*, 15(5), 555–568.

Cukier, K., Mayer-Schönberger, V., & de Vericourt, F. (2021). *Framers: Human advantage in an age of technology and turmoil*. New York: Dutton.

Ellen MacArthur Foundation. (2013). *Towards the circular economy: Economic and business rationale for an accelerated transition*. Cowes: Ellen MacArthur Foundation.

Faraj, S., Pachidi, S., & Sayegh, K. (2018). Working and organizing in the age of the learning algorithm. *Information and Organization*, 28(1), 62–70.

Flanagan, J. C. (1954). The critical incident technique. *Psychological Bulletin*, 51(4), 327–358.

Gioia, D. A., Corley, K. G., & Hamilton, A. L. (2013). Seeking qualitative rigor in inductive research: Notes on the Gioia methodology. *Organizational Research Methods*, 16(1), 15–31.

Grant, R. M. (1996). Toward a knowledge-based theory of the firm. *Strategic Management Journal*, 17(S2), 109–122.

Jarzabkowski, P. (2005). *Strategy as practice: An activity-based approach*. London: Sage.

Kogut, B., & Zander, U. (1992). Knowledge of the firm, combinative capabilities, and the replication of technology. *Organization Science*, 3(3), 383–397.

Lacy, P., & Rutqvist, J. (2015). *Waste to wealth: The circular economy advantage*. London: Palgrave Macmillan.

Lichtenthaler, U. (2019). An intelligence-based view of firm performance: Profiting from artificial intelligence. *Journal of Innovation Management*, 7(1), 7–20.

Nambisan, S., Lyytinen, K., Majchrzak, A., & Song, M. (2017). Digital innovation management: Reinventing innovation management research in a digital world. *MIS Quarterly*, 41(1), 223–238.

Nonaka, I., & Takeuchi, H. (1995). *The knowledge-creating company: How Japanese companies create the dynamics of innovation*. New York: Oxford University Press.

Orlikowski, W. J. (2002). Knowing in practice: Enacting a collective capability in distributed organizing. *Organization Science*, 13(3), 249–273.

Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic capabilities and strategic management. *Strategic Management Journal*, 18(7), 509–533.

Timmermans, S., & Tavory, I. (2012). Theory construction in qualitative research: From grounded theory to abductive analysis. *Sociological Theory*, 30(3), 167–186.

Vial, G. (2019). Understanding digital transformation: A review and a research agenda. *Journal of Strategic Information Systems*, 28(2), 118–144.

Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). Thousand Oaks, CA: Sage.

Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. New York: PublicAffairs.

Mapping Green and Digital Strategy Adoption in the Packaging Industry: an Empirical Knowledge-Based Analysis

Marco Scatto[1], Chiara Grosso[2], Alberto Petroni[3], Karen Venturini[1]
[1] University of San Marino, San Marino, R.S.M.
[2] Antonio Ruberti, University of Studies of Roma "La Sapienza", Rome, Italy
[3] University of Parma, Parma, Italy

Abstract

The packaging industry is currently undergoing a profound transformation driven by the Twin Transition, which combines environmental sustainability and digitalisation as key levers for industrial competitiveness and regulatory compliance. European policy frameworks, including the Ecodesign for Sustainable Products Regulation and the Circular Economy Action Plan, increasingly require firms to integrate eco-design principles, lifecycle thinking, and digital traceability tools into their production and decision-making processes. However, despite the strategic relevance of these developments, the actual level of adoption of green and digital strategies within the packaging sector remains uneven and insufficiently documented. This study aims to map the maturity level of green and digital strategy adoption in the packaging industry, adopting a knowledge-based perspective consistent with Eco-design and focus the study investigates how firms are operationalizing sustainability- and digital-related knowledge to transit their production processes towards a more green and sustainable strategies.
In the context of the transition toward a circular economy and industrial digitalization (Ghobakhloo et al., 2020), the packaging sector faces complex challenges related to environmental sustainability and technological innovation. Digital transformation, integrated with eco-design principles and traceability across the product lifecycle, emerges as a strategic lever to reduce environmental impacts and ensure regulatory compliance (Pagoropoulos et al., 2017; Kristoffersen et al., 2020). This study conducts an empirical, real-based analysis derived from a structured survey administered to entrepreneurs, scientist, consultant and operators in the packaging sector. The resulting mapping enables the identification criterias that allow to prioritize strategies for managing the end-of-life (EoL) phase of packaging materials, focusing on criteria such as environmental impact, economic feasibility, regulatory compliance, technological readiness, and stakeholder acceptance with Analytic Hierarchy Process (AHP) method derived from Multi-Criteria Decision Analysis (MCDA). The theoretical contribution lies in defining sustainability criteria for packaging in line with the new European regulations for Packaging and Packaging Waste. The expected results include a description of the state of the art on the twin transition and the structuring of sustainability criteria priorities among packaging producers and users in the management of end of life of packaging. This paper can help facilitate decision-making processes for packaging entrepreneurs in all the supply chain.

Keywords: Twin Transition, packaging industry, ecodesign, Multi-Criteria Decision Analysis, Analytic Hierarchy Process, survey

Paper type: Academic Research Paper

1 Introduction

In the context of the global transition toward circular economy (CE) models, the packaging sector represents a critical interface between industrial production systems and consumer behavior. Packaging has become a key target for systemic redesign in the shift to resource-efficient and sustainable production systems. Its prominent presence in consumer markets and significant contribution to post-consumer waste make it a pivotal intervention point for circular economy (CE) strategies (MacArthur, 2013). Regulations such as the European Union's Packaging and Packaging Waste Regulation (PPWR or 2025/40), effective from January 2025, impose stringent requirements whereby all packaging must be recyclable or reusable in technically and economically sustainable ways. However, the transition to circularity cannot be achieved solely through technological redesign; it also depends on active consumer involvement, whose preferences, awareness, and behavioral intentions determine the practical success of eco-designed solutions. To reduce the environmental impact associated with packaging, it is necessary to analyse production processes and materials involved, but particularly its end-of-life management. End-of-life packaging waste management has become a critical issue in pursuing sustainable development, especially in the European context, where environmental regulations and sustainability targets are stringent.

This paper presents:

- A systematised review to define the state of the art in the literature and thus the research questions and sustainability criterias that define decision making analysis.

- Application of AHP methodologies (Analytic Hierarchy Process (AHP) to construct a decision tree using the overall objective (strategy selection).

Following the bibliographic analysis, it emerged that neither the maturity level of packaging sector companies with respect to the Twin Transition nor the most relevant sustainability criteria to use in end-of-life packaging management are described in the literature. Moreover, there is a gap in business practices and a research gap in scientific publications on the application of AHP-based decision-making methods in end-of-life packaging design management. For this reason, this paper addresses the following two research questions (RQs):

- *RQ1: What is the maturity level of packaging sector companies with respect to the Twin Transition, i.e., the integration of green and digital strategies into their processes?*
- *RQ2: What are the sustainability criteria perceived as most relevant by companies in evaluating end-of-life packaging strategies?*

The work is developed to address these research questions. Current practices, policies, and technological innovations were examined in a sample of companies, professionals and academic researcher in the packaging sector.

Recent studies have highlighted that sustainable innovation is no longer purely a technological matter but a socio-technical process requiring coordination among various actors, including producers, designers, policymakers, and consumers (McDonough and Braungart, 2002). Co-design methodologies have proven effective in integrating stakeholder expertise and fostering cross-disciplinary mutual learning (Kirchherr, Reike, and Hekkert, 2017).

The main contribution of this study lies in its integrative analytical framework, which aims to bridge the intrinsic gap between company-driven design criteria and consumer-driven cognitive requirements regarding product sustainability. This comparative methodology was developed by mapping quantitative parameters derived from Multi-Criteria Decision Analysis (MCDA) such as AHP against qualitative insights from consumer behavior studies. In particular, the design criteria were selected through a literature review. Scientific articles were identified using the Scopus search engine with the keywords: circular packaging, technological innovation, customization end of life packaging, design for recycling, AHP method. The criterias have been submitted to a focus group of six experts from academy, association and industry for participatory discussion and to define those most appropriate for the packaging sector.

2 Theoretical Framework

The adoption of green and digital strategies in the packaging industry can be interpreted as a multidimensional and knowledge-intensive process, in which environmental, technological, regulatory, economic, behavioral, and strategic dimensions are deeply interconnected. Sustainable packaging does not simply derive from replacing one material with another; rather, it requires firms to evaluate the entire packaging system, from raw material sourcing and design to production, distribution, use, collection, recycling, recovery, or final disposal. In this perspective, life cycle thinking represents an essential theoretical foundation, as it allows for the identification of environmental impacts and trade-offs along the entire product system. Finnveden et al. (2009) highlight the importance of Life Cycle Assessment as a method to evaluate environmental impacts at different life cycle stages, while Laurent et al. (2014) show that, in waste management systems, environmental results are strongly influenced by system boundaries, treatment options, and methodological assumptions. This aspect is particularly relevant for packaging, where a solution that appears sustainable in one phase may generate negative effects in another. Jeswani, Krüger, and Russ (2021), for example, demonstrate that chemical recycling via pyrolysis, mechanical recycling, and energy recovery must be compared through a life cycle perspective, as none of these options can be considered ecologically less impactful regardless of the technological and systemic conditions in which it is applied.

Within this framework, the circular economy offers a broader conceptual perspective for interpreting the transition toward sustainable packaging. The shift from linear to circular models implies the redesign of packaging systems in order to reduce the use of virgin materials, increase recyclability, promote reuse, and close material loops. Hopewell, Dvorak, and Kosior (2009) show that plastic recycling represents both a great opportunity and a complex challenge, as its effectiveness depends on material properties, collection systems, contamination levels, sorting technologies, and the demand for secondary materials. Similarly, Lazarevic et al. (2010) highlight that plastic waste management must be evaluated from a life cycle perspective, as different recycling and recovery options entail uncertainty and environmental trade-offs. More generally, Geissdoerfer et al. (2017) define the circular economy as a new sustainability paradigm based on closing, slowing, and narrowing resource loops. Applied to the packaging industry, this paradigm suggests that sustainability cannot be achieved

solely through downstream recycling, but requires systemic redesign, supply chain coordination, and new forms of value creation.

However, the adoption of sustainable packaging solutions also depends on their technical and infrastructural feasibility. Ecologically preferable solutions can be implemented only when they are technologically mature, scalable, and compatible with production processes, logistics systems, and waste management infrastructure. Mankins (2009), through the concept of technology readiness assessment, provides a useful framework for evaluating whether a technology is sufficiently mature for industrial application. This is relevant for packaging innovations such as bio-based materials, biodegradable and compostable packaging, mechanical recycling, chemical recycling, reusable systems, and smart packaging, which may differ significantly in terms of maturity, reliability, and scalability. Furthermore, Yu et al. (2023) show that technological innovation toward sustainability is influenced not only by internal firm capabilities but also by external factors such as environmental regulation, public support, and financial development. This suggests that the feasibility of green packaging strategies depends on the interaction between internal knowledge, technological capabilities, and the broader institutional and infrastructural context.

The regulatory dimension further reinforces the need for a structured, knowledge-based approach. Packaging firms operate increasingly within a framework of environmental requirements concerning recyclability, recycled content, labeling, packaging minimization, waste prevention, and responsibility along the product life cycle. Weinrich et al. (2024), by analyzing decision-making processes regarding sustainable packaging options in the European food sector, show that packaging decisions are influenced by environmental, technical, economic, regulatory, and organizational considerations. In this perspective, the "Packaging and Packaging Waste Regulation – PPWR, Regulation (EU) 2025/40" assumes a central role, as it establishes a harmonized regulatory framework at the European level for packaging and packaging waste. The Regulation was published in the Official Journal of the European Union on January 22, 2025, entered into force on February 11, 2025, and will generally apply from August 12, 2026. It strengthens the centrality of requirements such as recyclability, minimum recycled content, reduction of superfluous packaging, waste prevention, reuse, refill systems, bio-based content, biodegradability and compostability of packaging, technical documentation, and harmonization of rules in the European internal market. Consequently, regulatory compliance no longer represents merely an administrative constraint, but becomes a driver of innovation, packaging redesign, and strategic adaptation. Firms must therefore not only understand regulatory obligations but also translate them into criteria for design, material selection, supplier management, reporting systems, traceability, and market communication.

Economic sustainability also plays a decisive role in the adoption of sustainable and digital strategies. Sustainable packaging solutions are more likely to be adopted when environmental improvements can be reconciled with cost efficiency, investment feasibility, and competitive value. Life cycle-based approaches can help firms identify environmental and economic trade-offs, for example, when higher material costs are offset by lower transportation, waste management, or regulatory compliance costs (Finnveden et al., 2009; Laurent et al., 2014). At the same time, circular economy strategies can generate value through resource efficiency, waste reduction, new business models, and increased resilience, although they often require investments in redesign, reverse logistics, and infrastructure (Geissdoerfer et al., 2017). Therefore, firms must evaluate sustainable packaging not only as an environmental or regulatory requirement but also as a potential source of long-term competitiveness.

Consumer behavior represents a further essential component of this transition. Packaging is a visible interface between firms and consumers, capable of influencing purchasing decisions, disposal practices, and trust in environmental claims. Testa, Iovino, and Iraldo (2020) show that information search mediates the relationship between the circular economy and consumer behavior in the purchase of circular packaging. This indicates that the acceptance of sustainable packaging by consumers depends on the availability, credibility, and clarity of information. Digital technologies can support this process through QR codes, smart labels, digital product passports, and traceability systems, which are capable of providing information on material composition, recyclability, environmental impact, and correct disposal. However, these tools are effective only if consumers perceive them as accessible and reliable.

Finally, sustainable packaging assumes strategic and reputational relevance. Packaging can strengthen brand positioning, stakeholder legitimacy, and competitive differentiation when environmental claims are credible and supported by evidence. Sharma, Soodan, and Chakraborty (2026) highlight the role of "green marketing" in promoting brand reputation, suggesting that sustainability-oriented communication can contribute to trust and competitive positioning. However, reputational value depends on the consistency between actual environmental performance and communication practices, as claims not supported by evidence can expose firms to accusations of greenwashing.

Overall, the theoretical framework suggests that the adoption of sustainable and digital strategies in the packaging industry must be interpreted as a knowledge-based process. Firms must integrate environmental knowledge derived from Life Cycle Assessment, circular economy knowledge related to material flows, technological knowledge regarding maturity and scalability, regulatory, economic, and organizational knowledge, consumer knowledge, and strategic communication knowledge. Digital technologies act as enabling mechanisms that support data collection, traceability, compliance monitoring, consumer communication, and strategic decision-making. In this sense, the adoption of sustainable packaging emerges as a complex process in which firms combine digital objectives and capabilities to redesign packaging systems, reduce environmental impacts, respond to regulatory pressures, create economic value, and strengthen market legitimacy.

3 Methodology

To construct the theoretical framework and identify the main dimensions influencing the adoption of green and digital strategies in the packaging industry in Europe market, this study adopted a systematized literature review conducted via the ScienceDirect database. The systematized review incorporates certain procedures typical of systematic reviews—such as explicit search criteria, transparent source selection, and structured synthesis of contributions—without claiming the same level of exhaustiveness as a full systematic review (Grant & Booth, 2009). The search focused on identifying contributions relevant to defining the study's main theoretical domains: Life Cycle Assessment, circular economy, plastic recycling, technological feasibility, European packaging regulation, consumer behavior, and green marketing. Priority was given to peer-reviewed articles published in journals indexed on ScienceDirect, selected based on their relevance to sustainable packaging, green and digital transitions, and the development of the empirical framework. These sources were complemented by institutional references to the Packaging and Packaging Waste Regulation (PPWR), Regulation (EU) 2025/40, as a central regulatory element for understanding the pressures driving the adoption of sustainable packaging solutions.

The literature synthesis was conducted narratively, in line with the heterogeneity of the selected contributions. The aim was not to aggregate quantitative results but to critically organize knowledge from diverse disciplinary fields and integrate it into a unified interpretive framework. Following Snyder (2019), the review thus served as a tool to develop the study's theoretical foundation, clarify key adoption factors, and highlight the role of digital technologies as enabling mechanisms. In this perspective, the systematized review provides the conceptual basis for the empirical analysis, while transparently acknowledging its primary limitation: it is structured and replicable in its general criteria but does not aim to exhaustively census all available literature on the topic.

The criterias defined in previous letterature analysis have been discussed with brainstroming method with 6 experts in packaging, from accademic, industry, associations, consulting company and finally used in AHP method in order to derive relative priorities. The Analytic Hierarchy Process (AHP) is a structured multi-criteria decision-making method that helps evaluate complex choices by breaking them into hierarchies of criteria, sub-criteria, and alternatives.

AHP, developed by Thomas L. Saaty in the 1970s, starts by modeling the decision problem as a goal-oriented hierarchy, with the top level as the main objective (e.g., optimal end-of-life strategy), middle levels as criteria like cost or environmental impact, and the bottom as options like recycling or composting. Experts then perform pairwise comparisons using a 1-9 scale (1 for equal importance, 9 for extreme preference) to derive relative priorities, forming judgment matrices whose eigenvalues yield normalized weights and eigenvector-based scores. A consistency ratio check (ideally under 0.1) ensures reliable judgments by measuring inconsistency against random indices.

AHP Key Steps are:

- Hierarchy construction: Decompose the problem into levels for tangible (e.g., costs) and intangible (e.g., sustainability) factors.
- Pairwise comparisons: Compare elements at each level to build matrices.
- Priority calculation: Compute weights via principal eigenvector; aggregate across levels for final rankings.
- Synthesis and sensitivity: Combine results and test robustness to input changes.

AHP excels in packaging waste management by integrating diverse factors like recyclability, biodegradability, regulatory compliance, and economic viability,critical for packaging materials. It handles stakeholder trade-offs. (e.g., environmental vs. cost priorities) in scenarios like filler selection or disposal optimization, providing transparent, defensible decisions amid EU regulations on microplastics and circular economy goals. AHP's flexibility with qualitative/quantitative data, consistency verification, and group decision support make it ideal for non-structured problems like end-of-life packaging, where multiple actors (producers, regulators) must

balance sustainability and feasibility. AHP will not replace the decision-making process itself but it will generate information needed to make the decision and will present it in a structured way. The AHP is a method of solving problems that involve prioritisation of alternative solutions and is based on the evaluation of a set of criteria.

4 Analysis Results

Decision-making techniques have been long dominated by the cost-benefit analysis that uses the criterion of maximum total net benefit to choose the optimum solution (Colorni et al., 1998). The main limitation of such an approach is that all the effects of an activity are evaluated in monetary terms, so that when aspects that cannot be quantified into monetary units (like social and environmental issues) are dominant, the method is completely inappropriate.

In the packaging sector, the final stage of the life cycle of a product, retirement, is one that has been often ignored. The retirement stage begins when the customer no longer wants the product. What happens to a product after the end of its usage is of growing importance to the producer because of consumer demands and other growing pressures such as new legislation that mandates recovery and recycling of some products by their manufacturers. Actually in the industry we have almost 4 alternative in management the end of life of packaging: reuse, mechanical recycling, organic recycling and chemical recycling.

In the AHP multicriteria analysis method used in this work, a hierarchical structure is formed using the overall objective (Selction of the optimal packaging end of life mangement strategy) as the root of the decision tree, with each main criterion placed as its child. Each criterion, in turn, is detailed to provide additional descendants. At the lowest level of the tree are the end-of-life options (reuse, mechanical, chemical, and organic recycling) assigned to each leaf node of the tree. We have created this decision tree in order to allow to the decision-maker performs pairwise comparisons at each level between criteria that share the same parent node and assigns them relative weights. In the next step after Fuzzification and defuzzification (Lee et al. 2010) is possible to convert fuzzy-set information into digital data evaluating all end-of-life options against each other in relation to each criterion, assigning them relative weights. All these comparisons lead to priority vectors that are propagated through the hierarchy to arrive at a final priority vector for the set of end-of-life options. The decision-maker is thus enabled to base their decision on various criteria to which they have themselves assigned weights.

The hierarchical structure of the criteria used in the AHP model in our study is represented in the following Figure, where the factors influencing the end-of-life management of packaging and the strategic alternatives considered are illustrated in decision tree.

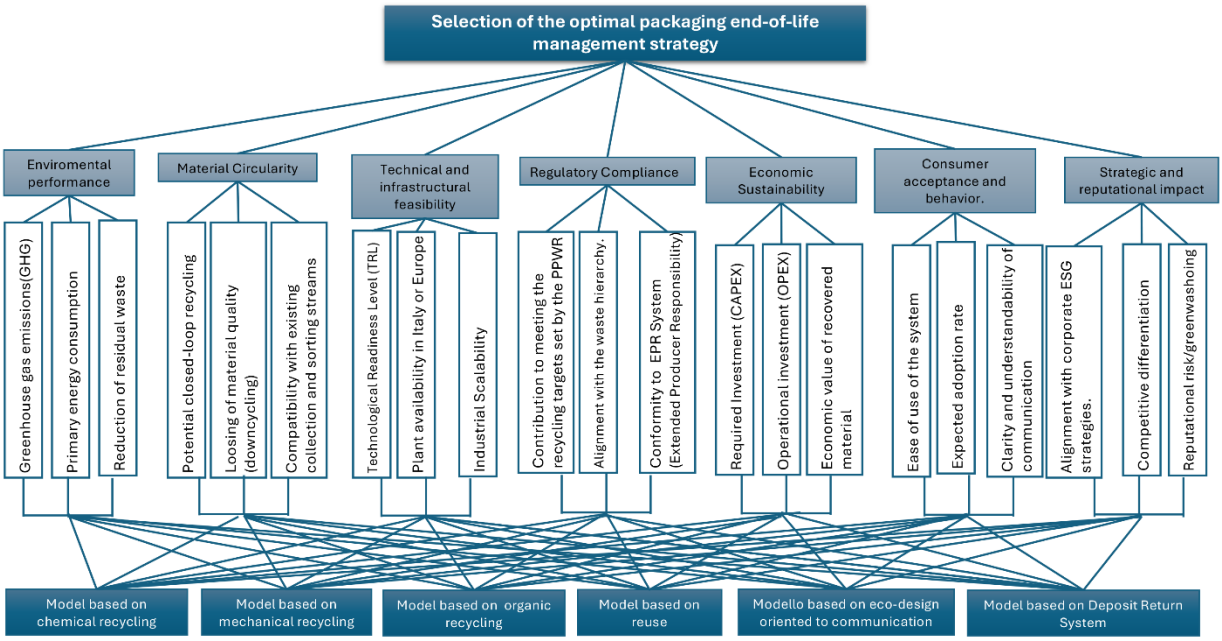


Figure 1: The hierarchical structure used by AHP in decision-making for end-of-life options for packaging: decision tree.

5 Conclusions

Decision-making for end-of-life of products, especially in packaging sector, is a difficult task. Many factors must be considered such as environmental, economic, social and so it becomes a multi- criteria decision problem. We have proposed one methodology, the AHP, that may help integrating all environmental, economic and social criteria in the decision-making process. The AHP is a simple and powerful tool that helps decision-makers structure and evaluate different alternatives. Incorporating environmental considerations, constraints stated by legislation and the voice of community in the decision-making process along with the economic judgements alters the decision for end-of-life of products. The main advantage of the AHP from the decision-maker point of view is that it s directly involved in the process. The decision-maker evaluates each pair of items he/she assesses. The result of the assessment is based only on the judgement of the decision-maker.In this paper we have demonstrate in the decision tree the criteria choosed and their connections.

6 Ethics Declatration

Our institutions does not require ethics approval for reporting individual cases or case series.

7 AI Declatration

No AI tools were used for creation of this paper.

References

Colorni et al., 1998, Decision support systems for environmental impact assessment of transport infrastructures, 30 April 1998, pp. 1-11.

Ellen Macarthur Foundation (2013), Report Towards the circular economy.

Finnveden et al. (2009) Recent developments in Life Cycle Assessment, Vol.91, October 2009, pp.1-21.

Geissdoerfer et al. (2017), The Circular Economy – A new sustainability paradigm?, Vol.143, 1 February 2017, pp.757-768.

Ghobakhloo et al. (2020) Industry 4.0, digitization, and opportunities for sustainability, Vol.252, 10 April, 119869.

Grant et al. (2009), A typology of reviews: an analysis of 14 review types and associated methodologies, Vol.26, 27 May 2009, pp.91-108.

Hopewell et al. (2009), Plastics recycling: challenges and opportunities, Vol. 364, 27 July 2009, pp.2115-2126.

Jeswani et al. (2021), Life cycle environmental impacts of chemical recycling via pyrolysis of mixed plastic waste in comparison with mechanical recycling and energy recovery, Vol.769, 15 May 2021, pp.144483.

Kirchherr et al. (2017) Conceptualizing the circular economy: An analysis of 114 definitions, Vol.127, December 2017, pp.221-232.

Kristoffersen et al. (2020) The smart circular economy: A digital-enabled circular strategies framework for manufacturing companies, Vol.120, November 2020, pp241-261.

Laurent et al. (2014) Review of LCA studies of solid waste management systems--part II: methodological guidance for a better practice, Vol.34, March 2014, pp. 589-6606.

Lazarevic et al. (2010), Plastic waste management in the context of a European recycling society: Comparing results and uncertainties in a life cycle perspective, Vol.55, Decembre 2010, pp.246-259.

Lee et al. 2010, Fuzzy systems to process ECG and EEG signals for quantification of the mental workload, Vol.142, May 2002, pp. 23-25.

Mankins (2009), Technology readiness and risk assessments: A new approach, Vol.65, November-December 2009, pp-1208-1215.

McDonough et al. (2002) Design for the Triple Top Line: New Tools for Sustainable Commerce, Vol.9, August 2002, pp.251-258.

Pagoropoulos et al. (2017) The Emergent Role of Digital Technologies in the Circular Economy: A Review, Vol.64, pp19-24.

Sharma et al.(2026), The Economics of Sustainability: Role of Green Marketing in Promoting Brand Reputation, 28 January 2026, pp.

Snyder et al.(2019), Literature review as a research methodology: An overview and guidelines, Vol.104, November 2019, pp.333-339.

Testa et al.(2020), The circular economy and consumer behaviour: The mediating role of information seeking in buying circular packaging, Vol. 29, Decembre 2020, pp. 3435-3448.

Weinrich et al. (2024), Decision-making processes on sustainable packaging options in the European food sector, Vol.434, 1 January 2024, pp. 13991.

Yu et al. (2023), Impacts of digital transformation on eco-innovation and sustainable performance: Evidence from Chinese manufacturing companies, Vol. 393, 20 March 2023, pp. 136278.

When Traditional Knowledge Meets Digitalization: The Italian High-Quality Olive Oil Value Chain and The Platformized Markets

Sergio Conterjnic, Cristina Simone, Antonio Laudando
Sapienza, University of Rome. Rome, Italy

Abstract

The Agrifood sector is one of the recognized pillars of the Italian economy, generating 31,8 % of its GDP (Federalimentare, 2023) and providing labor to 17% of the Italian labor force (Istat, 2017). It is also one of the engines of the Italian exports, generating 8,1 % of annual Italian exports (Federalimentare, 2023). Some products are considered symbols of the agrifood "Made in Italy", such as pasta, olive oil, cheese, ham and wines. The agrifood sector has performed very well in the last decades, representing an excellent example of Italian food culture worldwide. In this paper, we will focus on the high-quality olive oil value chain. In 1960 the Italian government established the controlled appellation system for olive oils, designing the Protected Designation of Origin (PDO or DOP) and Protected Geographical Indicator (PGI or IGP). The support for this classification and certification is that the olive oil producers operating in a particular area owe their success primarily to the quality of their products and the inseparable link between these products and their *terroir*, i.e. the particular geographical area and environment that give their olive trees and olive oil their distinctive features (Vaudour, 2002; Conti et al., 2024). In addition to this classification, the production and designation of "Extra Virgin Olive Oil" was controlled and protected by European regulations.

This research will analyze the high-quality olive oil value chain, to identify its internal and external strengths and weaknesses. In particular, the aim is to provide an accurate analysis of the strategic successful factors of the high-quality olive oil value chain, as well as its challenges related to market forces, climate change and certification and quality processes and identification. Secondly, this research aims to explore the implications of the new digital technologies applied to the agribusiness sector, in particular, the high-quality olive oil sector. The paper focuses on the pivotal role of digital business platforms in enabling high-quality olive oil producers to expand their market reach, optimize supply chain efficiency, and respond dynamically to evolving consumer preferences.

Concerning scholars, the paper integrates the extant literature on high-quality olive oil, clearly identifying the sources of advantages of PDO and PGI olive oils in digital platformized markets. Concerning practitioners, the paper provides a multidimensional strategic framework of analysis to support the main stakeholders within the high-quality olive oil both in physical and digital ecosystems in adopting sustainable and regenerative practices for a higher co-creation of value in the value chain. Eventually, the paper will provide suggestions for future research paths.

Keywords: Agricultural Value Chains, Agrifood, High-Quality Olive Oil, Digital Platforms.

Paper type: Academic Research Paper

1 Introduction

The Agrifood sector is one of the recognized pillars of the Italian economy, generating 31,8 % of its GDP (Federalimentare, 2023) and providing labor to 17% of the Italian labor force (Istat, 2017). It is also one of the engines of the Italian exports, generating 8,1 % of annual Italian exports (Federalimentare, 2023). Some products are considered symbols of the agrifood "Made in Italy", such as olive oil, pasta, cheese, ham and wines. The agrifood sector has performed very well in the last decades, representing an excellent example of Italian food culture worldwide. In this paper, we will focus on the high-quality olive oil value chain.

This research will analyze each of the stakeholders in one of the most important and recognized Italian agrifood value chains (Porter, 1985, 2001; Davis & Goldberg, 1957; German et al., 2020): the high-quality olive oil value chain, to identify its internal and external strengths and weaknesses. In particular, the aim is to provide an accurate analysis of the strategic successful factors of this value chain, as well as its challenges related to market forces, climate change and certification and quality processes and identification. Secondly, this research aims to explore the implications of the new customers' concerns and demands relative to the product and the care of the environment, both in the domestic and international markets.

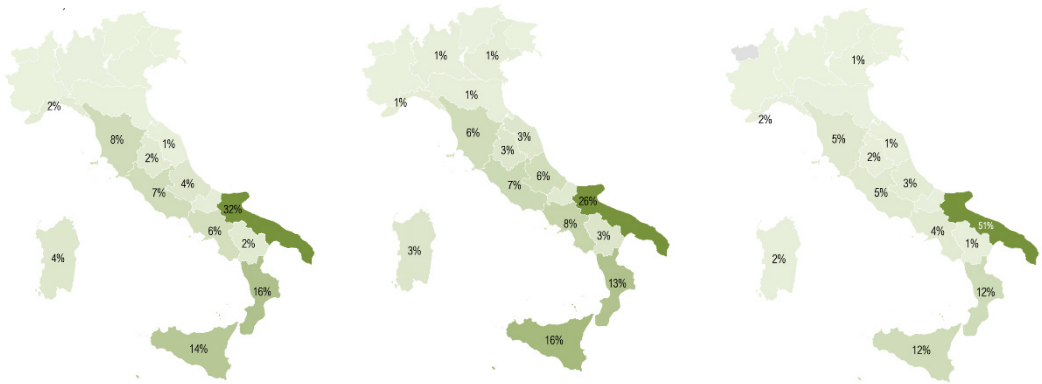
Regarding the first aim, a deep analysis of the vertical integration within the value chain will be performed. A cross-analyses of successful factors in the high-quality olive oil value chain will allow to discover the key strategic elements that made it to perform so well. Secondly, this research aims to explore on the pivotal role of digital business platforms in enabling high-quality olive oil producers to expand their market reach, optimize supply chain efficiency, and respond dynamically to evolving consumer preferences.

2 Italian Olive Oil Value Chain

Olive production and olive oil consumption was within the Italian heritage for millenniums, and it is one of the symbols of the agrifood "Made in Italy". Olive trees are planted all over the country, but for the Alps region. The total area is of 1,14 million hectares, mostly concentrated in the Southern regions: 62% – Puglia 32%, Calabria 16% and Sicily 14%.

The olive oil value chain consists of different stakeholders: (i) olive production (farmers), (ii) olive oil processing (mills, packing plants, refineries), (iii) marketing and distribution (supermarkets, small shops, gourmet food stores, exporters), (iv) consumers (domestic and international) and (v) government and services (consortiums, certification, etc.).

In Italy there are 619,000 olive producers and 4,240 olive mills. The pressing of the olives must be done within a day of harvesting, to avoid fermentation, so mills are in all the production areas. Producers are mostly small – average area of 1.8 hectares, 38% below 5 hectares – and mills are from small to medium capacity, mostly family or cooperatively run. Domestic consumption of olive oil is very high – 7.5 liters/person per year, due to the "Mediterranean diet", worldwide recognized for its health benefits.



(a) (b) (c)

Fig. 1: (a) Regional distribution of olive-growing area (average 2020-23), (b) Regional distribution of olive farms in 2020 and (c) Regional production (average 2020-23).

Source: Istituto di Servizi per il Mercato Agricolo-Alimentare (ISMEA) – Scheda del settore olio di oliva. Agosto 2025

Regarding mills, the larger producers have their own olive mill, sometimes with machinery for bottling and packing in different formats (cans or Tetra-Pack containers), while the small and medium producers bring their olives to an external mill, which can be managed privately or cooperatively. In both cases, those services are normally paid in kind, with a share of the oil. Small farmers use the oil for family consumption and for selling the extra to a network of relatives and clients. Some specialized bottlers buy bulk olive oil and bottle it with their own labels or with the private labels of large retailers.

Regarding the household domestic consumption, the two main retail channels are large retail chains (65%) and direct channels (farm and mill direct selling: 26%).

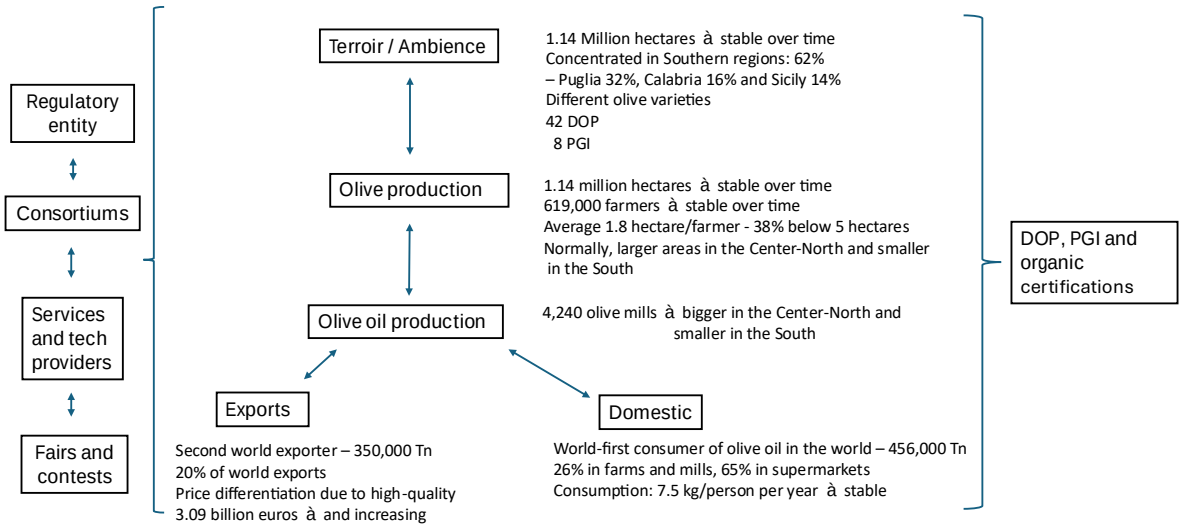


Fig. 2: Italian Olive Oil Value Chain.
Source: Self elaboration

Based on the average of the 4 campaigns 2021/22 to 2024/25, Italy is the world-first consumer of olive oil in the world – 456,000 Tn; second producer (after Spain) – 287,000 Tn; first importer – 524,000 Tn; and second exporter (after Spain) – 361,000 Tn, reaching 20% of world exports, but always with a price differentiation, due to quality recognition: the "Made in Italy" factor.

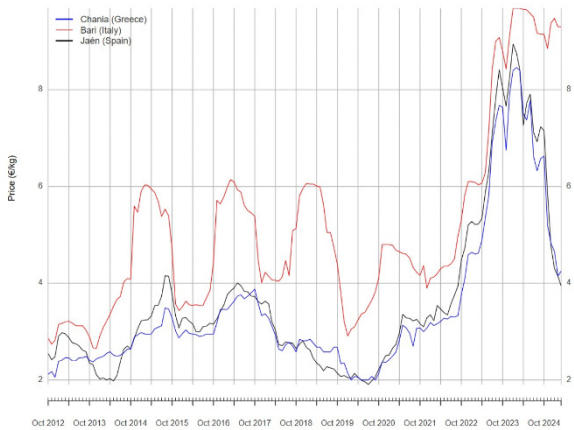


Fig. 3: Movements in producer prices (Oct 2012 / Mar 2025).
Source: International Olive Council: Olive sector statistics. March 2025

For the recognition of the different types and quality of olive oils and their identity, both in the domestic and international markets, the Italian government adopted in 1960 the controlled appellation system for olive oils, designing the Protected Designation of Origin (PDO or DOP) and Protected Geographical Indicator (PGI or IGP).

The effect of the PDO economy in Italian territories should not only consider the economic return of the productive sector, but to measure the value that "belongs" to those territories: the result of non-relocatable productions, expression of a collective heritage and a set of environmental, historical and cultural factors.

Regarding olive oil, there are 42 PDO and 8 PGI in Italy. These certifications apply only to 6.25% of the national production (13,000 Tn), while an impressive 24 % of the production is certified as "organic" – mostly due to the small area and natural production system, being this recognized by a price differentiation in the international markets for the Italian quality, in comparison with its competitors (Spain, Greece and Tunis). The quality and organic certification are important for customers in the international markets and in the big distribution

channels. As mentioned before, in Italy, a large volume of high-quality olive oil is sold directly from farmers and mills (26 %) showing that the closer the consumer to the producer, the less a certification is needed.

In the PDO-PGI market, a very interesting case is the region of Tuscany, that representing only 5% of national production is exporting 40% of the PGI olive oil, due to the international recognition of the region as high-quality food and wine provider.

In addition to this classification, the production and designation of "Extra Virgin Olive Oil" (EVOO) was controlled and protected by European regulations in the 'Art. 35 of EC Reg. 1513/2001 , which defines it as a product obtained from the first pressing of olives through mechanical processes, thus without the use of chemical processes or substances, under conditions that do not cause alterations in the oil and whose free acidity, expressed as oleic acid, is not more than 0.8 percent.

Organic production is to obtain foods and other agricultural products based on the use of processes with minimal impacts on the environment and the health of humans, plants and animals. Also, organic products are considered healthier and safer than conventional products, so it is not simply a matter of having a better taste and being more environmentally friendly. Income influences consumption of organic food, since are more expensive. The higher purchase of organically produced EVO oil is in gourmet food stores and, to a lesser extent, at farms gates or mills direct sales.

3 International Markets

In recent years, Italy has strengthened its role as the world's second-largest olive oil exporter, shipping in 2024/25 approximately 344,000 tons valued at over €3.09 billion. Major export markets are: the United States, Germany, France, Japan and Canada. The United States consume around 400,000 Tn/year (1.1 kg/person per year), but domestic production represents less than 2 per cent of consumption. The United States is the most important international customer, being Italy the larger supplier to the US (120,000 Tons, € 1billion): mostly Extra Virgin Olive Oil (89%), in small containers with private brands, reflecting quality recognition and price differentiation. This is a still-growing and high-price market for Italy.

4 The Platformization of the Italian High-Quality Olive Oil Value Chain

The growing adoption of digital business platforms is progressively reshaping the value chain of Italian high-quality olive oil, particularly with respect to traceability and commercialization (Parra-López et al., 2025; Muller et al., 2026). As a typical credence good, high-quality olive oil relies heavily on trust, origin, and quality attributes that are not directly observable by consumers (Sgroi et al., 2024). In this context, digital business platforms help reduce information asymmetries by integrating traceability systems, digital certifications, and product storytelling, thereby strengthening consumer trust and supporting price differentiation (Parra-López et al., 2025).

At the same time, digital platforms are becoming increasingly relevant for the commercialization of high-quality olive oil. While traditional channels still dominate the market, with large-scale retail accounting for approximately 65% of sales and direct channels for around 26% (ISMEA, 2025), a growing share of premium products is gradually shifting towards online and direct-to-consumer channels. This trend reflects broader dynamics in the Italian food sector: despite a recent stabilization, the combined value of online grocery and food delivery reached approximately €4.4 billion in 2023, underscoring the structural importance of digital channels for food consumption (Statista, 2025). Evidence from adjacent sectors, such as wine, highlights the magnitude of this transformation, with vertical platforms experiencing significant growth in recent years (Conterjnic et al., 2025). A similar, albeit less mature, dynamic can be observed in the olive oil sector, where vertical platforms are emerging. Examples include *Olive Oil Lovers* and *Citizens of Soil*, which focus on high-quality products and leverage digital channels to connect producers directly with consumers.

Although these platforms have not yet reached the scale observed in more mature sectors such as wine, they signal a clear process of platformization of the olive oil value chain, in which value creation increasingly relies not only on production and territorial embeddedness, but also on the ability to organize digital interactions, convey information, and orchestrate market access. In this evolving scenario, digital business platforms should therefore be interpreted not merely as additional sales channels, but as emerging infrastructural actors capable of reconfiguring coordination mechanisms, shaping consumer perception, and influencing the creation and appropriation of value along the value chain of Italian high-quality olive oil.

5 Conclusions

Extra Virgin Olive Oil (EVOO) is one of the most traditional Italian products, and its local production creates a virtuous circle of social and economic sustainability: supporting the viability of local economies, improving producers' incomes, and enhancing the quality of life within local territories. Furthermore, in recent decades, olive oil has also become one of the drivers of agrotourism, along with wine (Oil Routes, Towns of Oil, fairs and events), attracting tourists from all over the world into Italy's rural areas, generating income and labor opportunities.

However, the olive oil value chain is currently facing a set of structural challenges, including the impact of climate change, the spread of plant diseases, and the risk of declining production volumes (Fraga et al., 2020). These pressures are particularly critical in a sector characterized by high fragmentation, small-scale production, and strong territorial embeddedness (Martins et al., 2024).

In this evolving context, the findings of this paper suggest that digital transformation - particularly through the emergence of digital business platforms - is becoming an increasingly relevant dimension of competitiveness in the high-quality EVOO segment. Digital business platforms operate along two crucial and interrelated dimensions of value creation and appropriation. On the one hand, they enhance traceability and transparency by integrating digital certifications, data-driven information, and product storytelling, thereby strengthening trust mechanisms in a market where quality attributes are largely unobservable ex ante (Parra-López et al., 2025). On the other hand, they enable new forms of commercialization, allowing producers to access international markets more directly and enhance value capture through differentiation strategies (Conterjnic et al., 2025).

At the same time, this transformation contributes to a broader reconfiguration of the Italian olive oil value chain: from a predominantly linear, territorially bounded system to a more interconnected, platform-mediated ecosystem (Van Dijck et al., 2018; Poell et al., 2019; Barile et al., 2022). While this evolution creates new opportunities, particularly for small and medium-sized producers operating in the premium segment, it also introduces new structural tensions. In particular, the increasing reliance on digital business platforms tends to generate new forms of dependency, as market access, visibility, and competitive positioning become progressively mediated by digital intermediaries (Sauvagerd et al., 2024; Simone & Laudando, 2025; Simone et al., 2025).

Therefore, digital business platforms should not be interpreted merely as additional distribution channels, but as emerging infrastructural actors capable of reshaping coordination mechanisms, influencing consumer perception, and redefining value creation and appropriation within the olive oil sector (Podszun et al., 2021; Fletcher et al., 2024).

Building on these insights, future research could further investigate the role of digital business platforms in supporting quality recognition, traceability, and international positioning of high-quality Italian olive oil, as well as explore comparative analyses with other agri-food value chains, such as wine or cheese, to identify convergences and divergences in the process of platformization.

7 Ethics declaration

No ethical clearance is needed for the research referred to in this paper.

8 AI declaration

We did not use any AI tools for the research referred to in this paper.

References

Abbate, S., Centobelli, P., & Cerchione, R. (2023). The digital and sustainable transition of the agrifood sector. *Technological Forecasting and Social Change*, 187, 122222.

Barile, S., Saviano, M. & Simone, C. Service economy, knowledge, and the need for T-shaped innovators. *World Wide Web* 18, 1177–1197 (2015). https://doi.org/10.1007/s11280-014-0305-1

Barile, S., Simone, C., Iandolo, F., & Laudando, A. (2022). Platform-based innovation ecosystems: Entering new markets through holographic strategies. *Industrial Marketing Management*, 105, 467-477.

Bouhaddane, M., Samir, M. (2018). A Forecast of Internationalization Strategies for the Spanish Olive Oil Value Chain. *System Dynamics and Innovation in Food Networks, 2018.*

Brabazon, T., Winter, M., & Gandy, B. (2014). Digital wine: How QR codes facilitate new markets for small wine industries. Singapore: Springer.

Conterjnic, Sergio; Simone, Cristina; & Laudando, Antonio. Analyzing the Excellence in the Agrifood "Made in Italy": The Italian High-Quality Wine Value Chain and the Role of New Technologies. *20th International Forum on Knowledge Asset Dynamics*, Naples, Italy, 2025.

Conterjnic, Sergio; Simone, Cristina; & Laudando, Antonio. Developing a Multidimensional Framework to Sustain the Agribusiness Value Chains: The Case of Parmigiano-Reggiano Cheese. *19th International Forum on Knowledge Asset Dynamics*, Madrid, Spain, 2024.

Conti, M. E., Rapa, M., Simone, C., Calabrese, M., Bosco, G., Canepari, S., & Astolfi, M. L. (2024). From land to glass: An integrated approach for quality and traceability assessment of top Italian wines. *Food Control*, 158, 110226.

Davis, J. H., & Goldberg, R. A. (1957). *Concept of agribusiness*.

Federalimentare (2023). Rapporto Federalimentare-Censis: Il valore economico e sociale dell'industria alimentare italiana.

Fletcher, A., Crémer, J., Heidhues, P., Kimmelman, G., Monti, G., Podszun, R., ... & De Streel, A. (2024). The effective use of economics in the EU Digital Markets Act. *Journal of Competition Law & Economics*, 20(1-2), 1-19.

Fondazione Qualivita. Rapporto ISMEA-Qualivita 2025. Available at: https://www.qualivita.it/rapporto-ismea-qualivita-2025/#toggle-id-1

Forliano, C., De Bernardi, P., & Panero, M. (2024). The Covid-19 Impact on Food Digital Transformation: A Bibliometric Analysis of Academic and Practitioners' Perspectives. In Digital (Eco) *Systems and Societal Challenges* (pp. 29-53). Springer, Cham.

Fraga, H., Moriondo, M., Leolini, L., & Santos, J. A. (2020). Mediterranean olive orchards under climate change: A review of future impacts and adaptation strategies. *Agronomy*, 11(1), 56.

German, L. A., Bonanno, A. M., Foster, L. C., & Cotula, L. (2020). "Inclusive business" in agriculture: Evidence from the evolution of agricultural value chains. *World Development*, 134, 105018

Great Italian Food Trade. Italian olive oil: production, exports, and challenges. (2026). Available at: https://www.greatitalianfoodtrade.it/en/olive-oil/Italian-olive-oil-production-export-market/

International Olive Council (2026). Olive sector statistics – February/March 2026. Available at: https://www.internationaloliveoil.org/olive-sector-statistics-february-march-2026/#:~:text=At%20producer%20level%2C%20extra%20virgin%20olive%20oil,9%E2%80%9315%20February%2C%20%E2%82%AC440%20per%20100%20kg%20(+6%25)

ISMEA (2025). Scheda di Settore, Olio d'Oliva. Available at: https://www.ismeamercati.it/flex/files/1/6/f/D.cff78df6f341746db73f/Scheda_OLIO_agosto_2025.pdf

Istat (2017). *L'andamento della economia agricola*.

Martins, S., Pereira, S., Dinis, L. T., & Brito, C. (2024). Enhancing olive cultivation resilience: sustainable long-term and short-term adaptation strategies to alleviate climate change impacts. *Horticulturae*, 10(10), 1066.

Muller, L. F., Pereira, A. S., Santoyo, A. H., Becker, C., Herrmann, F. F., & Baierle, I. C. (2026). Digital Transformation Through Traceability: Enhancing Fraud Prevention and Economic Sustainability in the Olive Oil Industry. *Sustainability*, 18(3), 1475.

Parra-López, C., Abdallah, S. B., Hassoun, A., Jagtap, S., Garcia-Garcia, G., Hassen, T. B., ... & Carmona-Torres, C. (2025). Digital and Industry 4.0 technologies in olive farming and industry: Recent applications and future outlook. *Smart Agricultural Technology*, 101376.

Podszun, R., Bongartz, P., & Langenstein, S. (2021). The digital markets act: Moving from competition law to regulation for large gatekeepers. *J. Eur. Consumer & Mkt. L.*, 10, 60.

Poell, T., Nieborg, D., & Van Dijck, J. (2019). Platformisation. *Internet policy review*, 8(4), 1-13.

Porter, M. E. (1985). Competitive advantage: creating and sustaining superior performance. 1985. New York: *FreePress*, 43, 214.

Porter, M. E. (2001). The value chain and competitive advantage. *Understanding business processes*, 2, 50-66.

Sauvagerd, M., Mayer, M., & Hartmann, M. (2024). Digital platforms in the agricultural sector: Dynamics of oligopolistic platformisation. *Big Data & Society*, 11(4), 20539517241306365.

Sgroi, F., Sciortino, C., Giamporcaro, G., & Modica, F. (2024). Exploring the impact of beliefs and experiential factors on extra virgin olive oil consumption. *Journal of Agriculture and Food Research*, 15, 101056.

Simone, C., Laudando, A. (2025). Principles and obligations of the Digital Markets Act in regulating the economic power of gatekeepers: Positive, negative or trade-off effects? *Electronic Markets*, 35(1), 42.

Simone, C., Laudando, A., & Scardini, S. (2025). Digital platform ecosystems: A systemic view of power wielded by the platform orchestrators over complementors. *Industrial Marketing Management*, 131, 31-45.

Statista (2025). Online grocery market in Italy - statistics & facts. Available at: https://www.statista.com/topics/6497/online-grocery-market-in-italy/?srsltid=AfmBOoria6f6A2BWaM4p-9JEhMx88Jl-t67AOtbRgbeq46RNCE_IWIM

U.S. Department of Agriculture. Economic Research Service. (2024). Available at: https://www.ers.usda.gov/data-products/charts-of-note/chart-detail?chartId=110569

Van Dijck, J., Poell, T., & De Waal, M. (2018). The platform society: Public values in a connective world. *Oxford University Press.*

Vaudour E. (2002). The quality of grapes and wine in relation to geography: notions of terroir at various scales. *Journal of Wine Research*, 13 (2): 117-141.

Knowledge Management in Sustainable Energy Transitions: An Exploratory Agent-Based Approach

Domenico Camarda[1], Maria Rosaria Stufano Melone[2]
[1] Polytechnic University of Bari, Italy
[2] CNR-ISTC Research Associate, Trento, Italy

Abstract

Sustainable energy transitions pose governance challenges that cannot be addressed through technological optimization alone: they require the management of heterogeneous forms of knowledge — technical, local, tacit, and institutional — distributed among actors with diverging interests and evolving over time. This work proposes an exploratory approach based on Agent-Based Modeling (ABM) as a proof of concept for decision-support architectures oriented toward territorial knowledge management. The application context is the Apulia region, with specific reference to the municipality of Giovinazzo as a conceptual testbed.

The model, developed in NetLogo with a 150-month horizon, simulates interactions between strategic and operational agents through two cumulative systemic states (ecosystem health and territorial consensus), interpreted as proxies of tacit environmental knowledge and distributed institutional trust. Feedback flows among agents affecting approval probability represent the main source of non-linearity in the system.

Three policy scenarios are compared by varying the operational mode of the regulatory authority: permissive, balanced, and conservative. Scenario A (environmentally permissive) maximizes land conversion but leads to the collapse of both indices, with institutional sensemaking effectively blocked and delayed social response. Scenario B (balanced) keeps both systemic states above critical thresholds through a mechanism that can be interpreted as double-loop learning, without maximizing any single objective. Scenario C (environmentally conservative) preserves ecological and social conditions at the expense of energy coverage, illustrating a mirrored form of selective blindness. A structural sensitivity analysis on three parameters — administrative capacity of the municipal technical office, composition of the landowner base, and institutional permeability to social knowledge — shows that the latter exhibits a non-linear response with an identifiable optimal window, a result that directly informs the priorities of a subsequent participatory phase.

The main contribution is the proposal of a framework connecting organizational knowledge management with computational modeling, rather than offering predictive claims, and focusing instead on how the cognitive assumptions underlying decision-making can be made explicit and compared. The framework is intended to be transferable to other sustainability domains where success depends on the ability to govern distributed knowledge under conditions of complexity and uncertainty.

Keywords: Knowledge governance; Distributed knowledge; Energy transition planning; Agent-based modeling; Adaptive learning.

Paper type: Academic Research Paper

1 Introduction

The transition toward sustainable energy systems represents one of the most complex governance challenges of our time. In real territorial contexts, it cannot be reduced to a matter of technological efficiency alone; rather, it involves forms of knowledge that are technically specialized, locally embedded, and institutionally mediated, distributed among actors whose interests only partially overlap, and subject to change over time (Geels, 2011; Turnheim et al., 2015). The fundamental difficulty lies not in a lack of information, but in the multiplicity of interpretive frames through which such information is perceived, filtered, and translated into action.

Local authorities are required to evaluate energy projects whose effects propagate across interconnected domains — land, ecosystems, economy, communities — generating trade-offs that are not immediately visible and feedbacks that are often delayed (Head and Alford, 2015). In this context, knowledge is not a fixed input to decision-making, but a resource that is co-produced and continuously reshaped through interactions among actors (Nonaka and Takeuchi, 1995; Wenger, 1998).

This paper proposes an exploratory approach based on Agent-Based Modeling (ABM), conceived as a proof of concept for possible knowledge-oriented decision-support architectures. The aim is not to build a predictive model, but to explore how different architectures of knowledge representation may reveal — or alternatively conceal — critical dynamics of energy transitions. The application context is the Apulia region, with a specific focus on the municipality of Giovinazzo as a conceptual testing ground. The scientific contribution lies in conceptualizing ABM as a knowledge governance device: not only an analytical tool for complex systems, but

also an infrastructure that allows the cognitive assumptions of actors involved in distributed decision-making processes to become explicit, comparable, and open to negotiation.

The work is part of a broader research program addressing, in parallel, both the architectural dimension of decision-support systems — including the integration of ontologies, multi-agent simulation, and artificial intelligence components — and the epistemic dimension of knowledge governance. The two strands share the same simulation base but pursue distinct questions: the former investigates how to design a multi-level DSS for policy development; this paper focuses instead on how distributed knowledge is structured and transformed within such a system. Scenario results are therefore shared for methodological reasons, while the knowledge management perspective represents the original contribution of this article.

2 Theoretical Framework

2.1 Distributed knowledge and sociotechnical transitions

The literature on sociotechnical transitions has consistently shown that the shift toward sustainability involves deep processes of cognitive and institutional reconfiguration (Geels, 2011). The Multi-Level Perspective highlights how technological regimes are stabilized by cognitive routines embedded in actors' practices, and how transitions require the destabilization of such patterns through processes of collective learning (Rip and Kemp, 1998). In this sense, technological change cannot be separated from changes in the structure of shared knowledge.

Nonaka and Takeuchi (1995) conceptualize knowledge creation as a conversion between tacit and explicit dimensions through cycles of socialization, externalization, combination, and internalization (the SECI model). In complex organizations — and even more so in multi-actor governance systems — this conversion does not occur smoothly. Structural resistances, power asymmetries, and interpretive barriers tend to slow down or distort the process. Zollo and Winter (2002) show that organizational dynamic capabilities depend on the ability to articulate tacit knowledge embedded in operational routines, transforming it into explicit rules that can be revised at the strategic level.

In the context of energy transitions, this tension appears as a misalignment between the strategic intentions of governance authorities and the operational practices of territorial actors (Kern and Smith, 2008). Public decision-makers formulate objectives in explicit and aggregated terms, while local communities develop their evaluations through tacit and context-dependent knowledge — perceptions of landscape, memory of past impacts, trust (or distrust) in institutions — which are difficult to translate into formats compatible with formal planning.

2.2 Collective sensemaking and boundary objects

The concept of sensemaking (Weick, 1995) refers to a retrospective and social process through which actors construct meaning by assigning interpretive frames to available information. In energy transition contexts, this process is inherently conflictual: different actors activate different frames — economic development, environmental protection, territorial identity — which are not easily comparable and often lead to misunderstanding or opposition that cannot be resolved simply by providing more information.

Ostrom (2009) has shown that the governance of complex socio-ecological systems requires forms of knowledge co-production between institutional actors and local communities. Wicked problems, such as those encountered in territorial energy planning, are characterized by the absence of a single optimal solution and by the contextual and interdependent nature of objectives (Head and Alford, 2015). The issue is therefore not to identify the 'right' knowledge, but to build processes that allow its continuous integration across different interpretive frames.

Within this perspective, the concept of boundary object (Star and Griesemer, 1989) becomes particularly relevant. Shared analytical tools — flexible enough to adapt to multiple frames, yet robust enough to maintain coherence across communities of practice — can play a key mediating role in distributed governance processes. Computational models such as the one proposed here aim to perform this function, providing a shared reference through which actors with different forms of knowledge can interact and compare positions.

2.3 ABM as an epistemic device

Agent-Based Modeling has been increasingly used in the study of energy transitions because of its capacity to represent complex adaptive systems with heterogeneous actors, non-linear dynamics, and emergent behavior (Chappin and Dijkema, 2010; Ringler et al., 2016). Unlike aggregated optimization models, ABM allows

for the explicit representation of individual decision logics, local interactions, and systemic feedbacks, making collective outcomes not reducible to the sum of individual actions.

From a knowledge management perspective, ABM can be seen as an epistemic device: it makes explicit the assumptions about how knowledge is structured among actors — who knows what, how beliefs are updated, how ambiguous signals are interpreted — and allows the exploration of the consequences of different cognitive architectures (Moglia et al., 2018; Robinson et al., 2012). The model proposed here follows this line, with a specific focus on institutional knowledge governance, exploring how different configurations of decision rules, institutional capacity, and social feedback mechanisms generate qualitatively different system trajectories.

Differently from traditional participatory modelling — where stakeholders are directly involved in model construction (Voinov and Bousquet, 2010) — the present approach uses simulation as a device of ex-post explicitation. Algorithmically generated scenarios act as a kind of reflective surface against which stakeholders' tacit knowledge can be articulated in a later phase. The contribution is not to replace actors' judgment, but to provide a formal structure that makes implicit governance assumptions visible, comparable, and open to discussion.

3 Model Architecture

3.1 General structure

The model is developed in NetLogo (Wilensky and Rand, 2015) and simulates a time horizon of 150 months (approximately 12.5 years), with each tick corresponding to three months. The role of parameter values is to generate comparable dynamics across scenarios, not to reproduce real-world magnitudes.

The development of the code benefited from generative AI tools as cognitive support during the prototyping phase, rather than as autonomous components of the model. A central design choice is the distinction between active agents and cumulative systemic states: agents act according to their decision rules at each tick, while state variables evolve cumulatively, incorporating memory and path dependency — a key feature of systems in which knowledge is embedded in institutional routines (Zollo and Winter, 2002). Knowledge is therefore not represented as a centralized database, but as a distributed set of localized rules, selective perceptions, and institutionalized routines.

3.2 Strategic and operational agents

The distinction between strategic and operational agents represents the main theoretical device of the model. Strategic agents — the regulatory authority (AC) and photovoltaic firms (FVE) — can adapt their policies in response to evolving signals. Operational agents — the municipal technical office (UTC), routine landowners, and rigid environmental associations — tend to follow predefined procedures, introducing a form of cognitive inertia that makes the translation of strategic intentions into concrete actions more difficult (Zollo and Winter, 2002).

The model includes seven agent types:

- FVE (firms): proposing agents with offers ranging from 5,000 to 15,000 monetary units, oriented toward maximizing installed capacity.
- Landowners (OWN): either speculative (acceptance threshold reduced by 10%) or aware (considering environmental and community impacts). Their proportion is controlled by speculator-owns.
- AC (authority): approval probability varies by mode — permissive (0.70), balanced (0.50), restrictive (0.35), conservative (0.20) — and is modulated by social mobilization through alpha-social (0.20).
- UTC: limited administrative capacity, with a growing backlog acting as a procedural bottleneck.
- AGR (agricultural land): distinguishes between traditional and agrivoltaic installations; subject to possible desertification.
- IND (industry): energy demand expressed as satisfaction on a 0–100 scale.
- AMB (environmental groups): either rigid (systematic opposition) or selective (evaluation based on technology and context), reflecting ideological versus contextual knowledge.

3.3 Cumulative systemic states as proxies of collective knowledge

ESY and COM are aggregated modelling constructs interpreted here as proxies of collective knowledge — a heuristic reading rather than an empirical claim.

- ESY (Ecosystem State): index of ecosystem health (initial value 1.0, critical threshold 0.55). Impact per installation: 0.010/tick (traditional), 0.004/tick (agrivoltaic); recovery rate 0.01/tick. ESY represents tacit

environmental knowledge embedded in the territory: its erosion signals the emergence of consequences not anticipated in collective decision-making.

- COM (Territorial Consensus): index of territorial consensus (initial value 1.0, opposition threshold 0.60). Similar dynamics to ESY. COM represents distributed trust in the governance system: its decline signals a failure of institutional sensemaking.

The feedback between ESY and COM on approval probability is the main source of non-linearity: when both fall below threshold, the system shifts regime in a way that is difficult to reverse (Argote and Miron-Spektor, 2011).

Social mobilization is updated at each tick in response to three triggers (development pressure > 25%, share of traditional PV > 60%, active AMB > 50%) and reduces AC approval probability through alpha-social. The structural asymmetry between the growth rate of mobilization (0.05) and its spontaneous decay (0.02) encodes the path-dependent nature of trust erosion.

4 Results: Knowledge Dynamics in Simulated Scenarios

4.1 Scenario design and epistemic caution

Three policy scenarios are compared over a 150-month period by varying the operational mode of the regulatory authority (ac-mode). The numerical values reported below are outputs of the model under its current assumptions. They do not represent empirical estimates of the real Apulian context, nor are they derived from field data. Their meaning is relative and comparative: they describe differences between scenarios, rather than absolute levels. In this sense, they should be read as indicative orders of magnitude, useful for reasoning about directions of change rather than precise quantities.

Table 1**:** Comparison of policy scenarios (150-month simulation)

Indicator	A scenario (Permissive)	B scenario (Balanced)	C scenario (Conservative)	Critical threshold
Land conversion	~61%	~45%	~21%	—
Main technology	Tradit. PV (~70%)	Agrivolt. PV (~54%)	Agrivolt. PV (~90%)	—
ESY (ecosystem health)	≈ 0.42 ⚠	≈ 0.68 ✓	≈ 0.91 ✓	0.55
COM (people consensus)	≈ 0.38 ⚠	≈ 0.73 ✓	≈ 0.94 ✓	0.60
Social mobilization	High	Moderate	Low	—
Energy coverage	High	Moderate	~21% ⚠	—
UTC backlog	High	Manageable	Low	—

4.2 Scenario A, permissive policy: blocked institutional sensemaking

With a permissive AC mode (baseline approval probability 0.70), the model produces the highest land conversion (around 61%) and a dominance of traditional PV installations (about 70%). At the same time, both ESY (≈ 0.42) and COM (≈ 0.38) fall below their critical thresholds.

From a knowledge management perspective, the AC agent does not incorporate into its decision mechanisms the signals of deterioration accumulating in the system. Local tacit knowledge — expressed through selective perceptions of environmental groups and aware landowners — fails to enter the formal approval process in a meaningful way. The UTC backlog progressively saturates, creating a misalignment between strategic intent and operational capacity: what is decided at the top cannot be effectively translated into practice. This could represent a failure in tacit-to-explicit knowledge conversion (Nonaka and Takeuchi, 1995), an interpretation that would require empirical verification.

The high level of social mobilization observed in later stages emerges as a bottom-up adaptive response: the community gradually develops collective knowledge about the impacts of installations and translates it into political pressure. However, because collective sensemaking is inherently retrospective (Weick, 1995), this response occurs too late, when system degradation is already advanced.

4.3 Scenario B, balanced policy: dynamic equilibrium

With a balanced AC mode (baseline approval probability 0.50), the model generates moderate land conversion (around 45%), a more even technological mix (agrivoltaic ~54%), and keeps both ESY (≈ 0.68) and COM (≈ 0.73) above their thresholds. Social mobilization remains present but manageable.

In this configuration, the AC acts as a kind of cognitive filter, introducing friction into the approval process and indirectly favoring diversification. The ESY–COM feedback remains active and prevents the system from drifting toward collapse. Under the model's assumptions, this behavior can be interpreted as a form of double-loop learning (Argyris and Schön, 1978): the system does not simply correct errors within fixed rules, but adjusts those rules in response to feedback.

No single objective is maximized in this scenario. Instead, all dimensions remain within acceptable ranges. Within the model's assumptions, no single objective dominates; instead, all dimensions remain within acceptable ranges, suggesting a form of maintained balance.

4.4 Scenario C, conservative policy: inverse selective blindness

With a conservative AC mode (baseline approval probability 0.20), land conversion remains low (around 21%), agrivoltaic installations dominate (~90%), and both ESY (≈ 0.91) and COM (≈ 0.94) remain high. However, energy coverage is insufficient (around 21%).

This scenario represents a mirror image of Scenario A. Here, the system systematically privileges environmental and social knowledge, but does so at the expense of economic and technological considerations. The result is another form of cognitive failure — not due to excess permissiveness, but due to over-correction. What the model cannot determine — and what would require empirical investigation — is whether this over-correction reflects a genuine institutional pathology or a rational response to past permissiveness. The simulation encodes it as symmetric failure, yet reality may be more ambiguous.

4.5 Cognitive thresholds and irreversibility dynamics

The critical thresholds of ESY (0.55) and COM (0.60) act not just as indicators, but as behavioral discontinuities. When crossed, the system shifts into a different regime, with consequences that are difficult to reverse. This reflects the path-dependent nature of organizational learning (Argote and Miron-Spektor, 2011): losses in cognitive capital — both environmental knowledge embedded in the territory and institutional trust — do not recover in a linear way.

The comparison across scenarios suggests that preventive governance — maintaining conditions above thresholds — may be more effective than reactive governance, which intervenes only after degradation has occurred. This remains, however, a hypothesis that would require further validation in the participatory phase.

4.6 Structural sensitivity analysis

In the absence of multiple empirically calibrated runs, the qualitative robustness of the model is assessed by analytically deriving the expected directions of response when varying three parameters that are theoretically relevant for the KM framework: the administrative capacity of the UTC (κ), the proportion of speculative landowners (σ), and the weight of social mobilization on the approval probability of the AC (α).

Unlike a statistical sensitivity analysis based on multiple simulation runs, this approach analytically derives the expected directions of system response. The objective is not to measure output variance, but rather to verify the structural coherence of the model with the theoretical assumptions of the KM framework, and to identify which parameters produce qualitative regime shifts rather than simple scale variations.

4.6.1 Experiment 1 — UTC administrative capacity (κ)

With low κ (e.g. 1), the UTC backlog grows faster than it can be processed: fewer installations slow down the erosion of ESY and COM, but the accumulated development pressure fuels social mobilization independently. With high κ (e.g. 6), the flow is faster but more exposed to stochastic variability in approvals.

Expected direction: κ↓ → ESY/COM moderately higher but higher mobilization; κ↑ → faster convergence toward asymptotic values of the baseline scenario.

The qualitative structure of scenarios remains robust with respect to κ: the parameter regulates the speed of the trajectory, not its final regime. From a KM perspective, this highlights that UTC administrative capacity is an autonomous cognitive variable — not simply a multiplier of AC policy.

4.6.2 Experiment 2 — Composition of the landowner base (σ).

As σ increases, the acceptance rate of offers rises, as does the volume of applications entering the UTC backlog, with direct impact on ESY and COM. A relevant aspect is that, since the technological choice between traditional and agrivoltaic installations is determined by the AC rather than by landowners, σ does not affect the technological mix but influences overall volume.

Expected direction: σ↑ → greater erosion of ESY and COM, earlier mobilization; σ↓ → slower system, more favorable profiles.

The qualitative structure is robust: σ acts as an acceleration factor, not as a regime-shifting variable. From a KM perspective, this indicates that the cognitive composition of the landowner base represents an independent source of systemic risk.

4.6.3 Experiment 3 — Weight of social mobilization (α).

α modulates the reduction of approval probability as a function of active mobilization. With α = 0, the system becomes insensitive to social feedback: ESY and COM converge toward worse values because the self-regulation mechanism is deactivated. With α ≥ 0.35, mobilization acts as a brake that blocks decisions without proportionally improving the indices.

The response is non-linear, with an optimal window between 0.10 and 0.25 — neither deaf nor paralyzed — consistent with the notion of double-loop learning (Argyris and Schön, 1978).

α represents institutional permeability to social knowledge: its non-linearity is the most theoretically relevant result of the analysis, because it directly informs the priority questions of the participatory Phase 2.

Table 2**:** Structural sensitivity analysis (baseline scenario: ac-mode = balanced)

Parameter	**Levels explored**	**ESY/COM Expected direction**	**Unchanged qualitative structure?**	**Relevance for KM frameworks**
κ (UTC capacity)	1 / 3* / 6	κ↓ → Higher ESY/COM, higher mob.; κ↑ → faster convergence	Yes — κ regulates speed, not regime	Operational inertia as an autonomous KM variable
σ (speculator-owns)	20% / 50%* / 80%	σ↑ → greater ESY/COM erosion and earlier mob.; σ↓ → slower trajectory	Yes — acceleration factor, not regime change factor	Cognitive composition of micro-actors as systemic risk
α (alpha-social)	0 / 0.20* / 0.40	α=0 → Lower ESY/COM (deaf system); moderate α → preserves thresholds; high α → decision block	No — non-linear; optimal window 0.10–0.25	Institutional permeability: the most critical parameter of the framework

** = Baseline values for scenarios A/B/C. Directions derived analytically from NetLogo procedures; they do not constitute output from simulated runs.*

The comparison across the three experiments reveals an overall pattern consistent with the KM framework assumptions. For κ and σ, the qualitative structure of scenarios is robust: both parameters affect the speed of trajectories rather than their regime. For α, instead, the response is non-linear — institutional blindness at low values, decision paralysis at high values — with an identifiable optimal window. This represents the most relevant theoretical result of the structural analysis and directly guides the key questions for the participatory phase.

5 Implications for Research and Practice

5.1 Operational hypotheses for local authorities

The model formulates, within its own assumptions, three operational hypotheses that can be explored empirically in the participatory Phase 2. The fact that they are hypotheses rather than prescriptions does not make them less useful: structuring precise questions is already a contribution to knowledge governance.

UTC administrative capacity. Within the model assumptions, an under-dimensioned UTC does not simply produce procedural delays, but introduces a structural decoupling between the strategic intentions of the AC and the operational action of the institutional system. This decoupling erodes the coherence of public action and fuels development pressure even in the absence of new approvals. Investing in administrative capacity can therefore be interpreted, in KM terms, as investing in the system's ability to learn from its own decisions — transforming tacit knowledge embedded in bureaucratic practices into explicit and updatable operational competences.

Technological differentiation. Agrivoltaic installations generate structurally lower impacts on ESY and COM (0.004 vs 0.010/tick), contributing to the preservation of future approval conditions. Technological choice is therefore not neutral with respect to system trajectories: it produces long-term effects on institutional trust and ecosystem health that do not emerge from short-term analysis. The model suggests that incentivizing this choice through differentiated regulatory mechanisms may, from a systemic perspective, function as a strategy for preserving the cognitive and social capital of the territory.

Asymmetry of consensus. The structural asymmetry between mob-k-up (0.05) and mob-k-down (0.02) encodes in the model a dynamic widely recognized in the literature: trust builds slowly and is lost quickly. This implies that policies aimed at restoring consensus — once eroded — are structurally more costly than policies aimed at preserving it.

Precautionary governance approaches — building continuous structures of dialogue and knowledge co-production before tensions accumulate — may therefore be more effective than reactive approaches that intervene only when mobilization is already visible.

5.2 The model as a knowledge co-production infrastructure

The model is conceived as a potential boundary object (Star and Griesemer, 1989): a shared but flexible representation through which actors holding different forms of knowledge can reflect on assumptions, explore trade-offs, and compare alternative framings, without requiring full agreement on problem definitions or solutions. Its value does not lie in predictive accuracy, but in its capacity to structure complex conversations and make otherwise opaque dynamics visible.

This potential is articulated through a two-phase approach. Phase 1 — documented in this work — produces simulated scenarios that function as structured hypotheses: plausible configurations of the system, derived from explicit assumptions and comparable with one another. Phase 2, envisaged as a subsequent development, involves the critical confrontation of these scenarios with the tacit and contextual knowledge of local stakeholders — municipal authorities, agricultural operators, environmental associations, and communities — in structured participatory workshops.

Knowledge is not transferred from the model to stakeholders, but co-produced through their critical interaction (Argote and Miron-Spektor, 2011). This process also helps to reveal potential mismatches between model assumptions and real-world dynamics, making such divergences analytically useful rather than simply problematic.

5.3 Limits and future directions

The model has limitations that need to be explicitly acknowledged. The decision rules of agents are simplified compared to the real complexity of institutional and individual behaviors: in particular, the binary distinction between speculative and aware landowners does not capture the continuous range of real positions. Impact parameters (esy-impact-trad, com-impact-trad) are hypothetical-exploratory and require empirical calibration using real territorial data — GIS analysis on the Apulia region, cadastral data, and administrative data on the municipality of Giovinazzo.

The spatial dimension is treated in a simplified way, without accounting for landscape and infrastructural interdependencies. Finally, the model does not incorporate inter-period learning dynamics of institutional agents: AC and UTC operate with fixed parameters throughout the simulation, whereas in reality institutions may update their routines in response to accumulated experience.

Future directions include: the integration of ontological frameworks to formalize concepts, relations, and assumptions (Borgo et al., 2021), improving transparency and replicability; the introduction of adaptive learning capacities in institutional agents; and the extension of the framework to adjacent sustainability domains — water governance, forest planning, waste management — where similar problems of distributed knowledge governance arise under conditions of complexity and uncertainty.

6 Conclusions

This work has explored the use of Agent-Based Modeling (ABM) as a knowledge management tool in sustainable energy transitions, with a focus on the mechanisms through which distributed knowledge is represented, mediated, and transformed within a complex territorial governance system. The proposed model, applied in an exploratory way to the context of the Apulia region, makes it possible to render visible informational misalignments, cognitive thresholds, and alternative trajectories that would otherwise remain opaque to formal decision-making processes.

The main contribution is a framework that connects organizational knowledge management with computational modeling. This connection is not purely methodological, as it raises questions about how knowledge is represented and transformed across the two domains.

Within the assumptions of the proof of concept, the sustainability of the energy transition does not emerge from the optimization of single objectives, but from the maintenance of a dynamic balance between different and competing forms of knowledge: energy-related, environmental, social, and institutional. The critical

thresholds of ESY and COM identify discontinuities beyond which the system loses the capacity to learn from its past experience, entering regimes that are difficult to reverse.

The structural sensitivity analysis further indicates that institutional permeability to social knowledge (α) exhibits a non-linear response with an identifiable optimal window — a result that directly informs the priority questions for the participatory Phase 2 and represents the most theoretically relevant hypothesis of the framework.

Although grounded in the energy sector and in the Apulian context, the framework is intended to be transferable to other domains of sustainability in which success depends on the capacity to govern distributed knowledge under conditions of complexity and uncertainty. Sustainable transition cannot be reduced to a purely technical problem; it also involves a cognitive dimension, particularly where different forms of knowledge need to be articulated and negotiated over time.

7 Ethics declaration

This research is of a theoretical and computational nature and does not involve human subjects, personal data, field surveys, or direct interaction with communities or individuals. The municipality of Giovinazzo is used exclusively as a conceptual testbed for model parameterization; no empirical data concerning identifiable persons or organizations was collected or processed. Accordingly, no formal ethical clearance was required for the conduct of this study. The authors acknowledge, however, that the envisaged Phase 2 — involving participatory workshops with local stakeholders — will require appropriate ethical review prior to implementation, in accordance with applicable institutional and national guidelines.

8 AI declaration

Generative AI tools (specifically, large language model-based assistants) were used during the research process as cognitive support in two distinct phases: (1) during the prototyping of the NetLogo simulation code, where AI assistance contributed to the iterative drafting and debugging of agent procedures — as explicitly acknowledged in Section 3.1; and (2) during the drafting of portions of this manuscript, where AI tools were used to support textual editing, structural organization, and bibliographic consistency checks.

In both cases, all conceptual choices, theoretical framings, parameter definitions, scenario interpretations, and conclusions remain the sole responsibility of the authors. AI-generated content was critically reviewed, substantially revised, and integrated into an original argumentative structure developed independently by the authors. No AI tool was used as a source of scientific claims or bibliographic references.

References

Argote, L. e Miron-Spektor, E. (2011) "Organizational Learning: From Experience to Knowledge", Organization Science, Vol. 22, No. 5, pp. 1123-1137.

Argyris, C. e Schön, D.A. (1978) Organizational Learning: A Theory of Action Perspective, Addison-Wesley, Reading, MA.

Borgo, S., Borri, D., Camarda, D. e Stufano Melone, M.R. (2021) "An Ontological Analysis of Cities, Smart Cities and Their Components", in Nagenborg, M. et al. (a cura di), Technology and the City, Springer, Cham, pp. 365-387.

Chappin, E.J.L. e Dijkema, G.P.J. (2010) "On the Viability of the Dutch Electricity System in Transition", Infrastructure Complexity, Vol. 1, No. 1, pp. 1-12.

Geels, F.W. (2011) "The Multi-Level Perspective on Sustainability Transitions: Responses to Seven Criticisms", Environmental Innovation and Societal Transitions, Vol. 1, No. 1, pp. 24-40.

Head, B.W. e Alford, J. (2015) "Wicked Problems: Implications for Public Policy and Management", Administration & Society, Vol. 47, No. 6, pp. 711-739.

Kern, F. e Smith, A. (2008) "Restructuring Energy Systems for Sustainability? Energy Transition Policy in the Netherlands", Energy Policy, Vol. 36, No. 11, pp. 4093-4103.

Moglia, M., Cook, S. e McGregor, J. (2018) "A Review of Agent-Based Modelling of Technology Diffusion with Special Reference to Residential Energy Efficiency", Sustainable Cities and Society, Vol. 37, pp. 327-336.

Nonaka, I. e Takeuchi, H. (1995) The Knowledge-Creating Company, Oxford University Press, Oxford.

Ostrom, E. (2009) "A General Framework for Analyzing the Sustainability of Social-Ecological Systems", Science, Vol. 325, No. 5939, pp. 419-422.

Ringler, P., Keles, D. e Fichtner, W. (2016) "Agent-Based Modelling and Simulation of Smart Electricity Grids and Markets: A Literature Review", Renewable and Sustainable Energy Reviews, Vol. 57, pp. 205-215.

Rip, A. e Kemp, R. (1998) "Technological Change", in Rayner, S. e Malone, E.L. (a cura di), Human Choices and Climate Change, Battelle Press, Columbus, OH, pp. 327-399.

Robinson, D.T., Murray-Rust, D., Rieser, V., Milicic, V. e Rounsevell, M. (2012) "Modelling the Impacts of Land System Dynamics on Human Well-Being", Environmental Modelling & Software, Vol. 34, pp. 101-113.

Star, S.L. e Griesemer, J.R. (1989) "Institutional Ecology, 'Translations' and Boundary Objects", Social Studies of Science, Vol. 19, No. 3, pp. 387-420.
Turnheim, B., Berkhout, F., Geels, F., Hof, A., McMeekin, A., Nykvist, B. e Van Vuuren, D. (2015) "Evaluating Sustainability Transitions Pathways", Global Environmental Change, Vol. 35, pp. 239-253.
Voinov, A. e Bousquet, F. (2010) "Modelling with Stakeholders", Environmental Modelling & Software, Vol. 25, No. 11, pp. 1268-1281.
Weick, K.E. (1995) Sensemaking in Organizations, Sage Publications, Thousand Oaks, CA.
Wenger, E. (1998) Communities of Practice: Learning, Meaning, and Identity, Cambridge University Press, Cambridge.
Wilensky, U. e Rand, W. (2015) An Introduction to Agent-Based Modeling, MIT Press, Cambridge, MA.
Zollo, M. e Winter, S.G. (2002) "Deliberate Learning and the Evolution of Dynamic Capabilities", Organization Science, Vol. 13, No. 3, pp. 339-351.

Managing Circular Innovation Ecosystems through Knowledge Recombination: Assessing Ecosystem Health and ESG Outcomes in Phytoremediation-Based Circular Economy Models

Mercedes Grijalvo Martín[1], Joaquín Ordieres-Meré[1], Lorna Anguilano[2,3]
[1] Universidad Politécnica de Madrid, Department of Industrial Engineering, Madrid, Spain.
[2] Phyona Ltd., Stamford, UK
[3] Experimental Techniques Centre, Brunel University. Kingston Ln, Uxbridge UB8 3PH, UK.

Abstract:

The transition to a circular economy depends not only on technological innovation, but on the ability of innovation ecosystems to integrate diverse knowledge, coordinate interdependent actors, and deliver credible environmental, social, and governance (ESG) outcomes. Yet many circular initiatives remain fragmented and struggle to scale or demonstrate measurable impact. This paper addresses this challenge by advancing an ecosystem-level perspective grounded in knowledge management, focusing on how knowledge recombination enables effective circular innovation.

The study adopts a qualitative case study approach based on the Phyona/CSTO2NE use cases, which combine phytoremediation, phytomining, biomass valorisation, and green construction materials. In this ecosystem, contaminated land is reframed as a productive resource, where hyperaccumulator plants extract metals while generating biomass that can be further valorised into high-value applications. The case illustrates how environmental remediation, resource recovery, and material innovation are interconnected through complex knowledge flows spanning plant science, soil chemistry, materials engineering, and sustainability assessment.

The paper's main contribution is the development of an ecosystem health framework. Ecosystem health is defined as the capacity of a circular innovation ecosystem to sustain knowledge diversity, enable knowledge connectivity, coordinate actors, support adaptive learning, and ensure ESG credibility. These dimensions explain how circular systems evolve from isolated experimentation to scalable and legitimate solutions. ESG outcomes are conceptualised not as ex-post indicators, but as emergent properties of ecosystems capable of aligning environmental performance, social value, and governance over time.

The findings contribute to circular economy and knowledge management literature by positioning knowledge recombination as a core mechanism for system-level sustainability transitions. The framework also provides actionable guidance for managers and policymakers, highlighting the need to design circular initiatives as coordinated ecosystems supported by robust knowledge infrastructures and transparent assessment practices. By linking ecosystem health with ESG outcomes, the study offers a structured approach to evaluate and govern circular innovation in emerging sustainability domains.

Keywords: Circular economy; knowledge management; digital transformation; sustainability; ESG

Paper Type: Academic Research Paper

1 Introduction

The transition towards a circular economy requires more than replacing linear production processes with recycling or recovery practices. It depends on the capacity of innovation ecosystems to coordinate heterogeneous actors, integrate diverse knowledge domains, and transform environmental liabilities into sources of value. Circular economy models are therefore knowledge-intensive systems in which value emerges through the recombination of scientific, technological, organizational, and regulatory knowledge (Geissdoerfer et al., 2017; Scuotto et al., 2022).

Contaminated land and metal-rich waste streams represent both a challenge and an opportunity in this transition. Conventional remediation approaches treat contaminated soils as liabilities, while resource-intensive extraction processes continue to generate environmental impacts. Nature-based solutions such as phytoremediation and phytomining offer an alternative by enabling metal recovery from degraded land while contributing to ecological rehabilitation (Nti Nkrumah et al., 2016; van der Ent et al., 2015).

This paper examines a Phytomining-based circular ecosystem in which metal-rich biomass obtained from phytoecosystems grown on retired mines or other retired industrial sites, becomes a platform for multiple value pathways, including metal recovery, nanoparticle production, fertiliser generation, and potential applications in construction materials. These interconnected processes highlight the need to move beyond firm-level

innovation towards an ecosystem perspective, where value depends on coordinated knowledge flows and interdependencies among actors (Adner, 2017; Jacobides et al., 2018).

From a knowledge management perspective, such systems require the recombination of diverse knowledge bases, enabling the transformation of distributed expertise into scalable sustainability outcomes (Grant, 1996; Nonaka et al., 1996). However, existing ESG approaches often focus on outputs rather than the underlying conditions that enable long-term impact.

This paper addresses this gap by proposing ecosystem health as an intermediate construct linking knowledge recombination with ESG outcomes. Drawing on the Phyona/CSTO2NE case, it develops a framework to assess how circular innovation ecosystems generate credible and scalable sustainability value.

2 Theoretical Background

2.1 Circular economy as a knowledge-intensive transition

Circular economy represents a systemic transformation that requires coordination across actors, knowledge integration, and new business models (Kirchherr et al., 2017). Its effectiveness depends not only on material recovery, but on the ability to connect previously separated activities, resources, and value logics.

Sustainable business model literature shows that circular transitions require reconfiguration of value creation, delivery, and capture mechanisms (Bocken et al., 2014; Boons & Lüdeke-Freund, 2013).

For the purposes of this paper, circular economy is therefore approached as a knowledge-intensive transition. Its success depends on the ability to mobilize scientific, technical, organizational, regulatory, and market knowledge across ecosystem boundaries. This is particularly important in emerging circular economy fields, where technologies are still developing, markets are uncertain, and sustainability impacts must be demonstrated credibly. The Phyona/CSTO2NE use cases are consistent with this challenge because they connect phytomining, contaminated land regeneration, biomass valorisation, nanoparticle recovery, and green construction materials within a single circular innovation logic.

2.2 Knowledge management and knowledge recombination

Knowledge management provides a useful theoretical lens for understanding how circular economy ecosystems create value from heterogeneous resources. The knowledge-based view argues that knowledge is one of the most strategically significant resources of organizations, especially when it is specialized, distributed, and difficult to imitate (Grant, 1996). In complex innovation settings, value creation depends less on the possession of isolated knowledge assets and more on the ability to integrate knowledge held by different actors, disciplines, and organizations.

Nonaka et al. (1996) theory of organizational knowledge creation is also relevant because it emphasizes the dynamic interaction between tacit and explicit knowledge. Circular innovation depends on both codified knowledge, such as technical data, environmental assessments, standards, and experimental results, and tacit knowledge, such as field experience, stakeholder understanding, process know-how, and practical judgment. The interaction between these forms of knowledge enables organizations and ecosystems to generate new solutions under conditions of uncertainty.

In the context of circular economy ecosystems, knowledge recombination is especially important. Knowledge recombination refers to the process through which existing knowledge elements are connected in new ways to generate novel applications, processes, or value propositions. Circular economy initiatives often require such recombination because they seek to create value from materials, sites, or outputs previously treated as waste or externalities.

Absorptive capacity further explains how organizations identify, assimilate, and apply external knowledge (Cohen & Levinthal, 1990). This is important in circular innovation ecosystems because no single actor typically holds all the knowledge required to deliver a complete circular solution. The ecosystem's capacity to create value therefore depends on the ability to absorb and recombine knowledge across organizational boundaries.

From this perspective, knowledge management is not simply a support activity. It is a central orchestration mechanism through which circular economy initiatives become technically feasible, economically meaningful, and socially legitimate. In the case examined in this paper, knowledge management enables the translation of phytomining and phytoremediation knowledge into multiple circular pathways, including metal recovery, nanoparticle production, biomass valorisation, and material applications. The central theoretical issue is therefore not only whether the ecosystem possesses valuable knowledge, but whether it can recombine that knowledge into scalable and credible ESG outcomes.

2.3 Innovation ecosystems and interdependent value creation

Innovation ecosystem theory provides a complementary perspective by shifting attention from individual firms to systems of interdependent actors. Adner (2017) defines ecosystems as structures of alignment among partners that need to interact for a focal value proposition to materialize. This view is particularly useful for circular economy initiatives, where value creation often depends on the coordination of multiple actors who contribute different resources and capabilities. In contrast to linear value chains, ecosystems involve complementarities, mutual dependencies, and coordination challenges that cannot be fully managed through traditional hierarchical control.

(Jacobides et al., 2018) argue that ecosystems emerge when groups of organizations coordinate around complementarities that are not easily governed through markets or vertical integration. This is relevant to circular economy systems because they often require actors to collaborate around uncertain technologies, emerging standards, shared infrastructures, and evolving value propositions. The viability of the ecosystem depends on whether actors can align roles, manage interdependencies, and coordinate activities without necessarily belonging to a single organization.

These interdependencies mean that circular value creation cannot be attributed to any single actor. Instead, it emerges from the configuration of roles, relationships, and knowledge flows across the ecosystem. This is why an ecosystem perspective is appropriate for the present qualitative case study.

Dynamic capabilities theory also strengthens this argument. (Teece et al., 1997) define dynamic capabilities as the ability to integrate, build, and reconfigure internal and external competences in response to changing environments. Circular economy ecosystems require similar capabilities, but at an interorganizational level. They must sense opportunities in contaminated land and residual biomass, seize them through new technological and business model configurations, and reconfigure relationships as technical evidence, regulatory expectations, and market opportunities evolve. In this sense, ecosystem-level dynamic capabilities are essential for moving from isolated experimental activities to scalable circular economy models.

2.4 Ecosystem health as an intermediate construct

Although circular economy and innovation ecosystem literatures provide important insights, assessing the performance of circular innovation ecosystems remains challenging. Conventional performance indicators often focus on firm-level outputs, financial results, or environmental metrics. ESG frameworks, meanwhile, frequently emphasize reporting outcomes after activities have taken place. While such indicators are important, they may not fully capture whether the ecosystem has the underlying capacity to sustain circular value creation over time.

This paper therefore introduces ecosystem health as an intermediate construct between knowledge management capabilities and ESG outcomes. Ecosystem health refers to the capacity of an innovation ecosystem to maintain effective knowledge flows, coordinate interdependent actors, adapt to technological and regulatory change, and align value creation with environmental, social, and governance objectives. The concept is useful because it shifts attention from isolated outputs to the conditions that make long-term circular impact possible.

In the context of a circular innovation ecosystem, ecosystem health can be understood through several dimensions. First, knowledge diversity refers to the presence of multiple knowledge domains required to address complex sustainability challenges. Second, knowledge connectivity refers to the mechanisms through which actors exchange, translate, and integrate knowledge. Third, coordination capacity refers to the ability to align roles, responsibilities, incentives, and expectations across the ecosystem. Fourth, adaptability refers to the capacity to respond to new evidence, site-specific conditions, regulatory requirements, and technological constraints. Fifth, legitimacy refers to the ability to demonstrate credible environmental and social value to stakeholders.

These dimensions are particularly relevant to phytoremediation-based circular economy models. The effectiveness of such models depends on biological processes, environmental conditions, material properties, downstream applications, and stakeholder acceptance. If knowledge is fragmented, if actors are poorly coordinated, or if ESG claims cannot be substantiated, the circular model may remain experimental or fail to scale. Conversely, a healthy ecosystem can support learning, experimentation, validation, and diffusion across multiple use cases.

Ecosystem health also helps connect qualitative case study evidence with theory development. Rather than treating the case as a simple illustration of circular economy principles, the ecosystem health lens enables a structured analysis of how knowledge assets are mobilized, where coordination challenges arise, and how ESG outcomes become more or less credible. This supports theory-building because it identifies the mechanisms through which knowledge recombination contributes to circular economy performance.

2.5 ESG outcomes as emergent properties of circular ecosystems

ESG outcomes in circular ecosystems are better understood as emergent properties of knowledge, coordination, and governance processes rather than isolated reporting indicators. Environmental impacts relate to remediation effectiveness and resource recovery; social outcomes to land regeneration and stakeholder engagement; and governance to transparency and compliance. Their credibility depends on the ecosystem's capacity to integrate technical validation, environmental assessment, and stakeholder communication.

The relationship between circular economy and ESG is therefore mediated by ecosystem health. A circular initiative may have strong environmental potential, but weak governance or poor stakeholder engagement can undermine its legitimacy. Similarly, technical feasibility does not automatically translate into social value unless the ecosystem develops mechanisms for inclusion, communication, and benefit distribution. This is particularly important in cases involving contaminated land, recovered materials, and emerging nature-based technologies, where public trust and regulatory acceptance are essential.

Life cycle assessment and related sustainability assessment methods can strengthen ESG credibility by providing quantitative evidence of environmental impacts. However, such methods depend on data availability, methodological assumptions, allocation choices, and system boundaries. Therefore, they must be understood as part of a broader knowledge infrastructure rather than as standalone proof of sustainability. In circular ecosystems, ESG assessment becomes a learning mechanism that helps actors identify hotspots, compare alternatives, and improve the design of processes and value propositions over time.

In the case considered in this paper, ESG outcomes are linked to several interconnected pathways: contaminated land remediation, metal and nanoparticle recovery, fertiliser production, residual biomass valorisation, and potential applications in lower-carbon construction materials. These pathways require different forms of evidence and different stakeholder relationships. The credibility of the overall circular model therefore depends on the ecosystem's ability to integrate technical validation, environmental assessment, governance mechanisms, and stakeholder communication.

2.6 Implications for the qualitative case study

The theoretical background developed above supports the qualitative case study approach adopted in this paper. Case study research is appropriate when the aim is to understand complex, context-dependent phenomena in which boundaries between the phenomenon and its context are not clearly separable (Yin, 2018). It is also suitable for theory development when the research seeks to explain mechanisms, relationships, and emerging constructs rather than test predefined hypotheses (Eisenhardt, 1989).

The Phyona/CSTO2NE case is theoretically relevant because it combines several features that are central to circular innovation ecosystems. It involves a nature-based remediation and resource recovery model, the transformation of contaminated, or better actually, metal-rich biomass into multiple value streams, the potential integration of residual fibres into construction materials, and the need to substantiate ESG outcomes. It also requires collaboration across scientific, industrial, regulatory, and market domains. These features make it a suitable case for examining how knowledge recombination supports ecosystem health and ESG value creation.

Based on the literature reviewed, the paper adopts three guiding theoretical assumptions. First, circular economy initiatives should be analysed as systemic and knowledge-intensive transformations rather than as isolated technical solutions. Second, knowledge recombination across organizational and disciplinary boundaries is a core capability for circular innovation ecosystems. Third, ESG outcomes emerge from the health of the ecosystem, understood as its capacity to coordinate actors, sustain knowledge flows, adapt to uncertainty, and provide credible evidence of impact.

These assumptions guide the subsequent case analysis. The focus is not only on what the Phyona/CSTO2NE ecosystem produces, but on how knowledge assets are mobilized and recombined across the ecosystem to support circular value creation. In doing so, the paper contributes to knowledge management and circular economy literature by proposing ecosystem health as a conceptual bridge between knowledge-based capabilities and measurable ESG outcomes.

3 Research Design and Case Context

3.1 Research design

This study adopts a qualitative case study design to examine how knowledge recombination supports ecosystem health and ESG value creation in a phytoremediation-based circular economy ecosystem. A case

study approach is appropriate when the research objective is to understand a contemporary, context-dependent phenomenon in which the boundaries between the phenomenon and its setting are not clearly separable (Yin, 2018). It is also suitable for theory development where the aim is to identify mechanisms, relationships, and emergent constructs rather than to test predefined hypotheses (Eisenhardt, 1989; Siggelkow, 2007).

The case is treated as a revelatory and illustrative case because it brings together several features that are central to circular innovation ecosystems: nature-based remediation, critical material recovery, biomass valorisation, green construction applications, and sustainability assessment. Rather than analysing a single technology in isolation, the study examines the configuration of actors, knowledge assets, and value pathways that connect contaminated land, phytomining, bioprocessing, nanoparticle production, residual biomass, and cementitious material development. This ecosystem-oriented framing follows the logic of qualitative research in which complex empirical settings are used to refine conceptual understanding and generate theoretically meaningful insights (Dubois & Gadde, 2002; Gioia et al., 2013).

The empirical basis of the study consists of documentary and technical case material related to the Phyona/CSTO2NE use cases. The analysis focuses on identifying how knowledge flows across technical, organizational, environmental, and material domains, and how these flows contribute to the development of circular value propositions and ESG outcomes. The purpose is not to provide a full technical validation of each process, but to interpret the use cases as evidence of an emerging circular innovation ecosystem.

The analytical strategy follows an abductive logic, moving iteratively between theory and case evidence (Dubois & Gadde, 2002). The initial theoretical categories were derived from the literature on circular economy, knowledge management, innovation ecosystems, and ESG assessment. These categories were then refined through the case material, with particular attention to four dimensions: knowledge recombination, ecosystem coordination, circular value creation, and ESG credibility. This approach enables the paper to develop ecosystem health as a conceptual bridge between knowledge management capabilities and sustainability outcomes.

3.2 Case context: the Phyona/CSTO2NE circular innovation ecosystem

The focal case is centred on Phyona's phytoremediation- and phytomining-based circular economy model, commercialised by Phyona Ltd (www.phyona.co.uk) and further developed to integrate innovation for the decarbonisation of the construction sector in connection with CSTO2NE use cases. The extended abstract frames the case as a circular innovation ecosystem in which plant-based technologies are used to treat contaminated soils, extract heavy metals, and support the production of high-value nanoparticles. Rather than treating soil contamination as a residual environmental problem, the model reframes degraded land and metal-rich biomass as sources of regenerative value. This makes the case especially relevant for examining how circular economy initiatives can integrate remediation, resource recovery, and advanced material production within a broader socio-technical system.

The phytomining use case describes a workflow that integrates soil characterisation, plant selection, agronomic optimisation, biomass processing, and metallurgical recovery. Soil characterisation is used to assess contamination, metal bioavailability, and expected recovery potential, while plant selection and cultivation rely on hyperaccumulator species and agronomic practices that improve biomass yield and metal uptake. Downstream processing then converts metal-rich biomass into recoverable resources, although the use case also highlights biomass processing as a bottleneck for field-scale deployment because of the technical and economic challenges associated with transforming variable biomass into concentrated metal feedstock.

The environmental and ESG relevance of the case is reinforced by the life cycle assessment material included in the phytomining use case. Phyona Ltd commissioned an assessment covering three sites: Salford, Barnsley, and a hypothetical scaled future site. The study used a cradle-to-gate approach and considered climate change, water consumption, land use, and mineral resource scarcity. It also defined three functional outputs: one kilogram of filtered and dried nanoparticles, one kilogram of organic fertiliser, and one hectare of decontaminated land. This is important for the present paper because it shows that ESG outcomes are not treated as abstract claims, but as impacts that require system boundaries, allocation choices, functional units, and environmental indicators. Such use of life cycle thinking is consistent with established sustainability assessment practice, where environmental impacts must be assessed across defined product systems and life cycle stages (Huijbregts et al., 2016; ISO/TC 207/SC 5, 2006, 2022).

The CSTO2NE green cement use case extends the circular logic by examining how biomass and fibres can contribute to lower-carbon construction materials. The consortium has worked on two complementary approaches: bio-inspired composite design using bio-derived reinforcement materials and carbon-sink green mortars derived from phytomining biomass. The first approach explores natural fibres and carbonated matrices, while the second investigates residual fibres from the bioprocessing of metal-rich phytomining biomass. These

fibres are assessed for their reaction with mortar matrices and their potential use as reinforcement or functional additives that may improve workability, reduce Portland cement and water demand, and support lower embodied carbon construction materials.

The technical material also indicates that phytomining-derived fibres were characterised through chemical and microstructural methods and integrated into mortar formulations with varying water-to-cement ratios and fibre contents. The CSTO2NE methodology included sourcing biomass from hyperaccumulator plants, bioprocessing through Phyona's proprietary method for metal recovery, size fraction mapping of fibre residues, SEM/EDS, FTIR and TGA-PKM characterisation, mechanical testing under compression and flexure, and evaluation of thermal insulation potential. This reinforces the interpretation of the case as a knowledge-intensive ecosystem in which plant science, soil chemistry, bioprocessing, materials characterisation, construction technology, and sustainability assessment must be connected.

The case also involves a multi-actor configuration. The phytomining use case identifies collaboration among Phyona Ltd, Brunel University London, Silesian University of Technology, CPBI, and DTU, while the green cement use case consolidates contributions from Brunel University London, Phyona Ltd., University of Lleida, and Silesian University of Technology. These actors contribute different forms of expertise, including phytomining bioprocessing, fibre characterisation, mortar formulation, analytical testing, and market-oriented assessment. The case therefore provides a suitable empirical setting for analysing ecosystem health as the capacity to coordinate diverse actors and recombine distributed knowledge into credible circular and ESG outcomes.

Overall, the Phyona/CSTO2NE case is not examined as a finished commercial system, but as an emerging circular innovation ecosystem. Its analytical value lies in showing how environmental remediation, metal recovery, residual biomass valorisation, and construction material innovation can be connected through knowledge recombination. This makes it possible to explore how ecosystem health develops, where coordination challenges arise, and how ESG value can become more credible through technical validation, environmental assessment, and cross-sector collaboration.

4 Ecosystem Health Framework

Building on the Phyona/CSTO2NE case, this paper proposes ecosystem health as an intermediate construct linking knowledge recombination capabilities with credible ESG outcomes. Health refers to the capacity of the ecosystem to sustain knowledge flows, coordinate interdependent activities, adapt to uncertainty, and translate circular value creation into verifiable environmental, social, and governance outcomes.

The proposed framework is structured around five interrelated dimensions: knowledge diversity, knowledge connectivity, coordination capacity, adaptive learning, and ESG credibility.

First, knowledge diversity captures the breadth of expertise required to generate circular value. Phytomining-based circular models combine plant science, soil chemistry, microbiology, materials science, environmental assessment, construction technology, regulatory knowledge, and business model design.

Second, knowledge connectivity refers to the mechanisms through which distributed knowledge is exchanged, translated, and recombined across actors. Knowledge diversity alone does not produce ecosystem-level value unless actors are able to communicate across disciplinary and organizational boundaries. In the case analysed, laboratory characterization, soil intelligence, fibre testing, life cycle assessment, and material performance evaluation act as boundary objects that allow different partners to interpret and use each other's knowledge.

Third, coordination capacity concerns the ability of the ecosystem to align roles, responsibilities, incentives, and decision points. Circular innovation ecosystems are inherently interdependent: remediation activities affect biomass characteristics; bioprocessing affects metal recovery and residual fibre quality; material applications influence performance requirements; and ESG assessment depends on data generated throughout the system. Coordination capacity is therefore necessary to avoid fragmented innovation, where each actor optimizes its own contribution without strengthening the overall circular value proposition. This dimension is consistent with ecosystem theory, which emphasizes that value creation depends on the alignment of complementary actors around a shared value proposition (Adner, 2017; Jacobides et al., 2018).

Fourth, adaptive learning reflects the ecosystem's ability to respond to experimental evidence, site-specific conditions, regulatory constraints, and market feedback. Phytomining and biomass valorisation are not standardized processes; they depend on soil composition, plant performance, biomass variability, processing conditions, and downstream application requirements. The case material shows that different water-to-cement ratios, fibre sizes, and fibre concentrations affect the morphology and mechanical behaviour of cementitious formulations, while life cycle assessment identifies energy-related hotspots in the bioreagents preparation stage.

Finally, ESG credibility refers to the capacity to substantiate sustainability claims through transparent evidence and governance mechanisms. Circular economy initiatives can easily generate appealing narratives, but their legitimacy depends on whether environmental and social benefits can be demonstrated. In the Phyona/CSTO2NE case, ESG credibility is supported by the articulation of functional outputs, such as kilograms of nanoparticles, kilograms of organic fertiliser, and hectares of decontaminated land, as well as by environmental indicators including climate change, water consumption, land use, and mineral resource scarcity. Such assessment practices strengthen the connection between circular innovation and measurable impact.

Taken together, these five dimensions position ecosystem health as a practical and conceptual tool for assessing circular innovation ecosystems. A healthy ecosystem is not simply one that produces technically viable outputs, but one that can continuously recombine knowledge, coordinate actors, learn from evidence, and maintain the credibility of its ESG claims. This framework is especially relevant for emerging circular economy models, where technologies, markets, regulations, and stakeholder expectations are still evolving.

5 Discussion

The case analysis suggests that phytomining-based circular economy models should be understood as knowledge-intensive innovation ecosystems rather than as isolated environmental technologies. The Phyona/CSTO2NE case demonstrates that value creation emerges through the recombination of multiple knowledge domains: ecological knowledge concerning contaminated soils and hyperaccumulator plants; technical knowledge related to bioprocessing, metal recovery, and nanoparticle production; materials knowledge concerning residual fibres and cementitious matrices; and sustainability knowledge related to life cycle assessment and ESG reporting.

This finding contributes to knowledge management literature by showing that knowledge recombination is not only a driver of product or process innovation, but also a mechanism for circular value creation. In this case, circularity depends on the ability to reinterpret contaminated land and residual biomass as knowledge-enabled resources. What counts as “waste” or “contamination” changes when actors can connect plant-based remediation, metal recovery, biomass processing, and construction material development within a coherent ecosystem. Knowledge management therefore acts as an orchestration capability that enables the ecosystem to convert dispersed expertise into new circular value pathways.

The paper also contributes to circular economy literature by emphasizing the role of ecosystem health as a condition for credible circular impact. Much circular economy research focuses on material loops, business models, or environmental indicators. These are important, but they do not fully explain how circular systems become scalable, legitimate, and adaptive. The proposed framework suggests that circular economy initiatives require a healthy ecosystem capable of sustaining knowledge diversity, connectivity, coordination, adaptive learning, and ESG credibility. This perspective helps explain why promising circular technologies may fail to scale when governance, knowledge integration, or stakeholder alignment is weak.

Figure 1 illustrates the proposed ecosystem health framework. Knowledge recombination acts as the enabling mechanism through which diverse actors and knowledge bases are integrated. Ecosystem health is reflected in five dimensions (knowledge diversity, knowledge connectivity, coordination capacity, adaptive learning, and ESG credibility) which together support environmental, social, and governance outcomes in phytoremediation-based circular innovation ecosystems.

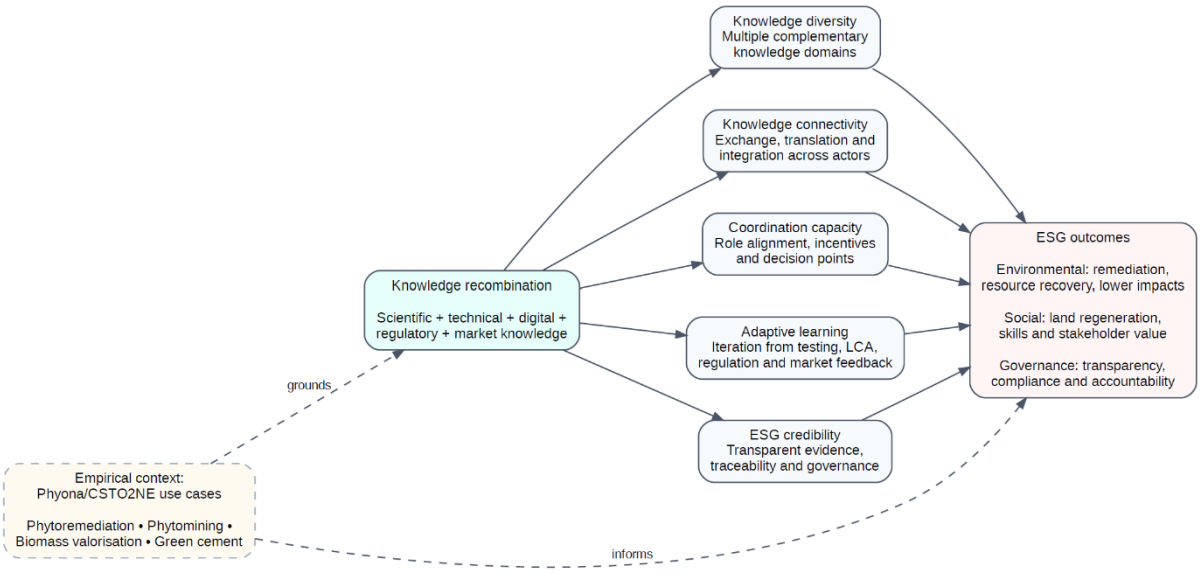


Figure 1. Ecosystem health framework for phytoremediation-based circular innovation

For managers, the framework highlights the importance of designing circular initiatives as ecosystems from the beginning. This means identifying not only technical partners, but also the knowledge interfaces required to connect them. Managers should pay attention to how data, test results, environmental assessments, and operational learning move across organizational boundaries. They should also treat ESG assessment as a learning process rather than as a final reporting exercise. In emerging circular systems, life cycle assessment, material testing, and site-specific monitoring can help identify hotspots, refine processes, and strengthen the credibility of sustainability claims.

For policymakers, the case illustrates the importance of supporting circular innovation ecosystems that connect environmental remediation with industrial valorisation. Contaminated land, mine residues, and marginal sites can be reframed as strategic assets when combined with appropriate knowledge, regulation, and market development. Policy support should therefore not focus only on individual technologies, but also on enabling infrastructures, testing facilities, data standards, certification pathways, and cross-sector partnerships. This is particularly relevant for critical raw materials, low-carbon construction, and nature-based solutions, where public legitimacy and regulatory confidence are essential.

The case also has implications for ESG practice. ESG outcomes in circular ecosystems should not be treated as isolated indicators attached to a finished product. They emerge from the way the ecosystem is organized and governed. Environmental outcomes depend on remediation effectiveness, resource recovery, carbon impacts, and material substitution. Social outcomes depend on land regeneration, local engagement, skill development, and the creation of new opportunities around degraded sites. Governance outcomes depend on transparency, traceability, role clarity, and credible evidence. Ecosystem health therefore provides a useful bridge between operational circularity and ESG performance.

However, the study also has limitations. The analysis is based on qualitative documentary and technical case material rather than longitudinal observation or primary interviews with all ecosystem actors. The case should therefore be interpreted as theory-building and exploratory rather than as a basis for statistical generalization. In addition, several technical pathways remain under development, particularly regarding the large-scale performance of phytomining-derived fibres in cementitious materials and the commercial scalability of biomass processing. Future research could extend this work through longitudinal case studies, comparative ecosystem analysis, stakeholder interviews, and quantitative assessment of ecosystem health indicators.

6 Conclusions

This paper examined how knowledge recombination supports ecosystem health and ESG value creation in a phytoremediation-based circular economy ecosystem. Grounded in the Phyona/CSTO2NE use cases, the paper argued that circular innovation should not be analysed only as a technical process of material recovery, but as an ecosystem-level phenomenon requiring the coordination of heterogeneous actors, knowledge domains, and value pathways.

The central contribution is the proposal of an ecosystem health framework for circular innovation. The framework identifies five dimensions: knowledge diversity, knowledge connectivity, coordination capacity,

adaptive learning, and ESG credibility. These dimensions explain how circular economy ecosystems can move from fragmented experimentation towards scalable and legitimate impact. In the case analysed, contaminated land, hyperaccumulator plants, recovered metals, nanoparticles, residual biomass, and cementitious materials are connected through knowledge flows that enable new forms of circular value creation.

The paper contributes to knowledge management research by positioning knowledge recombination as a core capability for circular innovation ecosystems. It contributes to circular economy literature by showing that ecosystem health can act as an intermediate construct linking knowledge management practices with ESG outcomes. It also offers practical guidance for managers and policymakers seeking to design, govern, and assess circular economy initiatives that combine nature-based solutions, resource recovery, and industrial applications.

Overall, the study suggests that credible circular economy transitions require more than technical feasibility. They require healthy ecosystems capable of learning, coordinating, adapting, and demonstrating impact over time.

7 Ethics Declaration

Ethical clearance was not required for this research, as the study is based on documentary analysis, technical use-case material, and non-human-subject research. No personal data, interviews with human participants, or sensitive individual-level information were collected or analysed.

8 AI Declaration

Generative AI tools were used to improve the linguistic flow and grammatical accuracy of this manuscript, as the authors are non-native English speakers. However, the final content has been thoroughly reviewed and edited by the authors, who take full responsibility for the integrity and accuracy of the published work.

References

Adner, R. (2017). Ecosystem as Structure. *Journal of Management*, *43*(1), 39–58. https://doi.org/10.1177/0149206316678451

Bocken, N. M. P., Short, S. W., Rana, P., & Evans, S. (2014). A literature and practice review to develop sustainable business model archetypes. *Journal of Cleaner Production*, *65*, 42–56. https://doi.org/10.1016/J.JCLEPRO.2013.11.039

Boons, F., & Lüdeke-Freund, F. (2013). Business models for sustainable innovation: state-of-the-art and steps towards a research agenda. *Journal of Cleaner Production*, *45*, 9–19. https://doi.org/10.1016/J.JCLEPRO.2012.07.007

Dubois, A., & Gadde, L. E. (2002). Systematic combining: an abductive approach to case research. *Journal of Business Research*, *55*(7), 553–560. https://doi.org/10.1016/S0148-2963(00)00195-8

Eisenhardt, K. M. (1989). Building theories from case study research. *Academy of Management Review*, *14*(4), 532–550.

Geissdoerfer, M., Savaget, P., Bocken, N. M. P., & Hultink, E. J. (2017). The Circular Economy – A new sustainability paradigm? *Journal of Cleaner Production*, *143*, 757–768. https://doi.org/10.1016/J.JCLEPRO.2016.12.048

Gioia, D. A., Corley, K. G., & Hamilton, A. L. (2013). Seeking Qualitative Rigor in Inductive Research. *Organizational Research Methods*, *16*(1), 15–31. https://doi.org/10.1177/1094428112452151

Grant, R. M. (1996). Toward a knowledge-based theory of the firm. *Strategic Management Journal*, *17*(SUPPL. WINTER), 109–122. https://doi.org/10.1002/SMJ.4250171110;CSUBTYPE:STRING:SPECIAL;PAGE:STRING:ARTICLE/CHAPTER

Huijbregts, M. A. J., Steinmann, Z. J. N., Elshout, P. M. F., Stam, G., Verones, F., Vieira, M., Zijp, M., Hollander, A., & van Zelm, R. (2016). ReCiPe2016: a harmonised life cycle impact assessment method at midpoint and endpoint level. *The International Journal of Life Cycle Assessment 2016 22:2*, *22*(2), 138–147. https://doi.org/10.1007/S11367-016-1246-Y

ISO/TC 207/SC 5. (2006). *ISO 14040:2006 - Environmental management — Life cycle assessment — Principles and framework*. ISO Standards. https://www.iso.org/standard/37456.html

ISO/TC 207/SC 5. (2022). *ISO/TS 14074:2022 - Environmental management — Life cycle assessment — Principles, requirements and guidelines for normalization, weighting and interpretation*. ISO Standards. https://www.iso.org/standard/61117.html

Jacobides, M. G., Cennamo, C., & Gawer, A. (2018). Towards a theory of ecosystems. *Strategic Management Journal*, *39*(8), 2255–2276. https://doi.org/10.1002/SMJ.2904;WGROUP:STRING:PUBLICATION

Kirchherr, J., Reike, D., & Hekkert, M. (2017). Conceptualizing the circular economy: An analysis of 114 definitions. *Resources, Conservation and Recycling*, *127*, 221–232. https://doi.org/10.1016/J.RESCONREC.2017.09.005

Nonaka, L., Takeuchi, H., & Umemoto, K. (1996). A theory of organizational knowledge creation. *International Journal of Technology Management*, *11*(7–8), 833–845.

Nti Nkrumah, P., M Baker, A. J., Chaney, R. L., Erskine, P. D., Echevarria, G., Louis Morel, J., van der Ent, A., Nkrumah, P. N., M Baker, A. J., Erskine, P. D., van der Ent, A., Chaney, R. L., Echevarria, G., & Morel, J. L. (2016). Current status and challenges in developing nickel phytomining: an agronomic perspective. *Plant and Soil 2016 406:1*, *406*(1), 55–69. https://doi.org/10.1007/S11104-016-2859-4

Scuotto, V., Le Loarne Lemaire, S., Magni, D., & Maalaoui, A. (2022). Extending knowledge-based view: Future trends of corporate social entrepreneurship to fight the gig economy challenges. *Journal of Business Research*, *139*, 1111–1122. https://doi.org/10.1016/J.JBUSRES.2021.10.060

Siggelkow, N. (2007). Persuasion With Case Studies. *Https://Doi.Org/10.5465/Amj.2007.24160882*, *50*(1), 20–24. https://doi.org/10.5465/AMJ.2007.24160882

Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic Capabilities and Strategic Management. *Strategic Management Journal*, *18*, 509–533. https://doi.org/10.1002/(SICI)1097-0266(199708)18:7

van der Ent, A., Baker, A. J. M., Reeves, R. D., Pollard, A. J., & Schat, H. (2015). Commentary: Toward a more physiologically and evolutionarily relevant definition of metal hyperaccumulation in plants. *Frontiers in Plant Science*, *6*(JULY), 1–3. https://doi.org/10.3389/FPLS.2015.00554/TEXT

Yin, R. K. (2018). *Case study research and applications* (Vol. 6). Sage Thousand Oaks, CA.

Economic Potential of ESCO Financing Models in Public Lighting: Evidence from Hungary

Nikolett Deutsch[1], László Berényi[2]
[1] Corvinus University of Budapest, Budapest, Hungary
[2] University of Miskolc, Miskolc, Hungary

Abstract

The technical capacity of LED street-lighting retrofits to achieve energy savings and reduce operating and maintenance costs is well established in the literature. This study contributes to the field of municipal energy-efficiency financing by presenting a two-sided discounted cash-flow analysis of public lighting modernization, comparing direct public financing with the ESCO-based model. Unlike most existing studies, which focus on either technical optimization or financing structures, this research models municipal and ESCO cash flows separately and examines their interaction within a unified financial framework for Hungarian municipalities. This issue is highly relevant in Hungary, where street lighting constitutes a major operational expense for municipalities with limited investment capacity and debt constraints. Regulatory and ownership limitations, lack of experience, uncertainty regarding realized savings, and weak procurement capacity further impede municipal retrofit activities. The analysis demonstrates that financial attractiveness depends not only on energy-saving potential as well as on tariff structures, contract duration, and contractor fees. The findings indicate that energy-efficient LED retrofits may remain financially unattractive for municipalities under direct self-financing, whereas the ESCO model can improve municipal outcomes by shifting upfront investment burdens, even if it reduces the provider's return. The results show that key contractual parameters primarily redistribute economic value rather than alter the underlying savings potential, which helps explain the frequent emergence of conflicts of interest in ESCO contracts. By quantifying these trade-offs and identifying threshold values at which municipal and ESCO incentives diverge or align, this study improves understanding of when and under what conditions ESCO financing constitutes a viable solution for small Hungarian municipalities. These insights prove particularly relevant for policy design and contract structuring in contexts characterized by budget constraints, tariff uncertainty, and limited local investment capacity.

Keywords: Energy Service Company, business model innovation, financial analysis, sustainability.

Paper type: Practical Paper

1 Introduction

Public lighting refers to the continuous, regular, fixed location artificial illumination of public spaces, such as roads, squares, parks, and private areas open to public traffic, operated according to a predefined schedule. Its primary purpose is to ensure traffic safety and protect public and private property at night. Municipalities are responsible for providing safe, reliable, lawful, and affordable public lighting services. At the same time, they are increasingly expected to improve energy efficiency, reduce greenhouse gas emissions, and mitigate environmental impacts such as light pollution. Thus, municipalities face a multi-objective challenge in public lighting management, and often struggle to deliver these services consistently due to several financial, technical, organizational, and regulatory obstacles. As Mercado et al. (2018) stress, public lighting services account for 30-50% of total municipal electricity consumption, and their share of municipal budgets varies with settlement size, ranging from 5-10% in larger cities to as much as 20% in smaller locations. As budgetary constraints restrict investment capacity as well as strategic flexibility, and municipalities are particularly exposed to energy price shocks, they tend to favour short-term solutions that give immediate cost reductions (Polzin et al., 2018). While increasing energy prices may strengthen the economic grounds for efficiency investments, they can also render decision-making more reactive and politically constrained (Lockwood & Selwyn, 2011). Due to aging systems, higher O&M costs, limited local expertise, procurement issues, and stricter regulatory and data reporting requirements, modernizing existing infrastructure is challenging. Although LED technology and adaptive lighting systems can deliver notable energy savings (Tsankov et al., 2025), the relatively high investment costs can delay implementation (Doulos et al., 2019; Balázs et al., 2023). Some authors (Valentiny, 2008; Pintér et al., 2018) also call attention to the fact that municipalities must operate under strict constraints on borrowing capacity, ownership structures, and tariff regulation, and in many cases, public funding schemes are insufficient, poorly targeted, or poorly aligned with the specific conditions of local projects. Polzin et al. (2016b) emphasize that procurement, implementation, and long-term management often exceed local governance capacities, making it more difficult for municipalities to implement retrofit projects. Similarly, Plepys and Richter (2016) highlight that

knowledge gaps, fragmented responsibilities, and risk-averse interpretations of procurement regulations hinder the uptake of innovative lighting solutions.

Municipal street lighting retrofits can be financed in many ways. Novikova et al. (2017) summarize options for municipalities, highlighting their advantages and disadvantages. Self-financing relies on revenues or grants, sometimes supplemented by internal contracting or revolving funds to ease taxpayer burden. However, limited revenues or weak capacity hinder project development, leading local governments to seek external financing like debt, municipal bonds, or institutional investment, with repayments from taxes or energy savings. Private contractors support retrofit projects through contractual schemes, covering design, financing, and implementation. Large infrastructure projects may involve private ownership and operation under leasing or concession agreements, with ownership reverting later. Funding often uses project finance via special-purpose vehicles, or crowdfunding for community involvement. Nowadays, there is a shift toward third-party financing models such as PPPs, ESCO schemes, and EPCs, in which investments are repaid from energy savings, either guaranteed or shared.

According to Bertoldi and Rezessy (2005), energy service companies develop, implement, and finance energy-efficiency or renewable energy projects, with project costs, including investment expenditures, recovered from the income stream generated by energy savings or renewable energy production. Bertoldi et al. (2014, p. 5) describe ESCOs as entities that guarantee energy savings or provide the same service at lower costs, with remuneration tied to savings and the ability to finance or help arrange financing for energy systems. Municipalities transfer financial responsibility for infrastructure updates to ESCOs via contracts, which include fees for planning, financing, and implementation. Operational activities are managed separately (Wardhana, 2024). This arrangement allows off-balance-sheet investment and simplifies procurement by enabling authorities to hire specialized service providers.

Business models, or financing schemes used for public energy efficiency projects, are well described in the relevant literature. Several case studies (World Bank, 2015; Marques & Geddes, 2019; Kurniawan & Kurniawan, 2022; Patra et al., 2025) evaluate how these options operate in practice, examining financing methods, asset ownership, and risk-sharing arrangements. Within the EPC and ESCO literature, public lighting is often seen as a promising application area, as LED-based renovations are relatively simple in technical terms, their benefits are easy to communicate, and the investment logic is often more transparent than in other energy-efficiency applications (Novikova et al., 2017). However, the empirical literature on public lighting is predominantly case-based, which limits the comparability and generalizability of the results. Quantitative studies on street lighting modernization primarily use techno-economic optimization or multi-criteria decision-making approaches. These papers often focus on the comparative assessment of replacement strategies, technology selection, and investment alternatives (Carli et al., 2018; Hermoso-Orzáez et al., 2019; Juričić et al., 2020; Gordić et al., 2021). Although these approaches play a major role in improving technical efficiency and decision support, they mostly neglect detailed analysis of project-level financing structures, contractual pricing mechanisms, and financial risk and benefit-sharing between the parties. A further branch of the literature examines ESCO and EPC structures primarily from a legal, institutional, and accounting perspective, with particular attention to public procurement rules, public debt classification, and state aid constraints (Polzin et al., 2016a,b). This approach provides an important analytical framework for understanding the institutional and regulatory feasibility of projects, but it is not intended to model in detail the project-level cash flows, tariff structures, or profitability profiles associated with specific technological uses. However, surveys of large-scale Western European projects show that while reducing energy and maintenance costs improves economic prospects, overall profitability remains sensitive to investment and financing costs (Campisi et al., 2017; Campisi et al., 2018). Studies based on German and Scandinavian cases show that municipalities often use EPC to reduce their investment needs and financial risks. This is especially important for municipalities with limited internal capital, technical expertise, or project management capacity (Polzin et al., 2016a,b; Aasen et al., 2016; Jensen et al., 2013). Studies focusing on Central and Eastern European countries mainly highlight the barriers and risks of third-party financing, and place greater emphasis on procedural feasibility, procurement structures, and legal frameworks than on detailed financial analysis of projects (Mikuŝová et al., 2020; Wacinkiewicz & Słotwiński, 2023). Furthermore, some authors (Bobbino et al., 2014; Kurowska-Pysz & Kunikowski, 2021; Wacinkiewicz & Słotwiński, 2023) argue that EU regulations on EPC, public accounting, and state aid strongly affect whether ESCO contracts are recorded on municipal balance sheets. In addition, financial analysis in many studies focuses primarily on institutional and accounting compliance issues, while the actual project-level cash flow structures are mostly only partially or implicitly reflected. While numerous business models, financing, and contracts for public lighting modernization are documented, much of the literature focuses on case studies or institutional issues. So it can be stated that there is still a lack of transparent, generalizable analyses that model municipal and ESCO cash flows, tariffs, savings sharing, and transaction costs in a unified framework.

This paper explores the ESCO model's potential to boost energy efficiency and cost savings for Hungarian municipalities facing limited funds, debt, aging infrastructure, and a lack of expertise. A Hungary-specific model integrating street lighting retrofits with ESCO financing is proposed as a decision-support tool. The research develops a hypothetical business model and an Excel-based decision analysis to replace fixtures with LED luminaires, using indicators such as NPV, ROI, and IRR to evaluate feasibility from municipal and utility perspectives. Cost data and forecasts support legislative and engineering considerations. Sensitivity analyses examine how variables such as energy charges, grants, contractor fees, payments, and contract length affect savings for both parties.

The structure of the paper is as follows: Section 2 outlines the model development and main assumptions. Section 3 presents the results of the financial and sensitivity analyses. Section 4 discusses the key findings in relation to the relevant literature. The conclusion provides a summary and final remarks.

2 Research Design

Hungary has achieved a substantial reduction in energy intensity over the past decade, with a 23.28% decline since 2015, according to Eurostat data. In parallel with this trend, an increasing number of municipalities have begun upgrading obsolete public lighting infrastructure by replacing traditional luminaires with energy-efficient technologies. As illustrated in Figure 1, these modernization efforts contribute to a measurable reduction in the overall electricity consumption of public lighting. While these trends are positive, they also mask persistent structural challenges. Energy intensity in Hungary remains well above the EU average, suggesting considerable untapped energy efficiency potential, particularly in the public sector and municipal infrastructure. In these areas, old systems and limited investment capacity continue to constrain modernization efforts (European Commission, 2025). Hungarian municipalities face a layered challenge, as they are under strong pressure to reduce electricity expenditure, a burden that is especially pronounced in smaller settlements. Local Hungarian governments must manage aging infrastructure, budgetary constraints, a lack of expertise, and a regulatory environment that does not always correspond to technically optimal solutions (Vidovszky Németh & Schanda, 2011; OECD, 2023; Vasvári & Longauer, 2024; Gellén & Horváth, 2026). Pintér et al. (2018) demonstrate that for many small municipalities, the upfront investment costs of lighting retrofits are prohibitively high without external funding, rendering several modernization pathways financially unviable in the absence of non-repayable public support. Drawing on international experience, ESCO-type intermediation can offer a mechanism for delivering financing, technical expertise, and risk management to Hungarian municipalities.

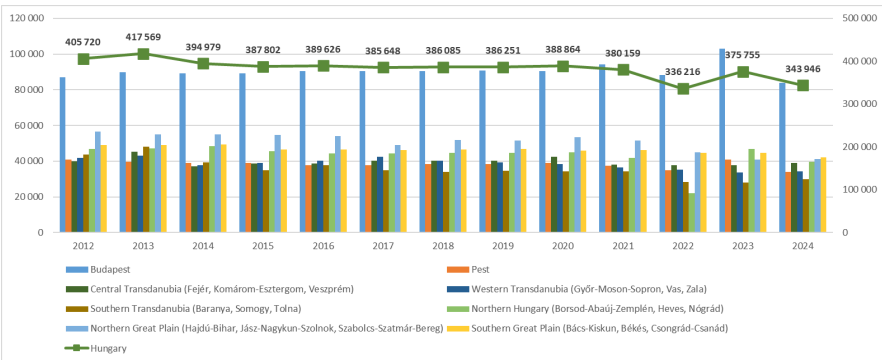


Figure 1. Energy consumption for public lighting - Regional distribution (in 1000 kWh)

This paper evaluates the financial viability of energy-efficiency investments using an ESCO financing model, focusing on modernizing municipal public lighting systems in Hungary. A hypothetical case study examines upgrading the public lighting system in a small Hungarian settlement with LED luminaires under an ESCO agreement. Based on an engineering review, it proposes a complete redesign using LED technology to reduce light pollution, improve lighting, and diminish energy use by 20%. Tables 1 and 2 show the technical details. The Excel-based decision model assumes that, under an ESCO scheme, the municipality's financial burden shifts to an energy service company via a contractual agreement. The service provider covers planning, financing, and retrofit costs, while lighting operation is managed under a separate contract. To assess the value of the ESCO model, the feasibility was evaluated for two options:

- Option 1 - Direct Investment by the municipality: The municipality, in collaboration with the energy company, modernizes its public lighting system by replacing existing luminaires with LED units. The full project cost is paid by the municipality in the year of the investment.

- Option 2 - ESCO-Based Model: The energy company handles the investment and operation. The municipality repays the cost over time in accordance with the ESCO financing agreement. The municipality repays the investment amount (including VAT) over 8 years. The initial payment at project start is zero, and from the second year, the municipality pays 1% of the investment amount annually as a contractor fee.

The common underlying assumptions of the feasibility analyses conducted for the options can be summarized as follows:

- Only luminaire replacement considered (no pole changes).
- Project lead time: 1 year.
- No investment subsidy or loans are assumed.
- Investment cost: 52 675 EUR.
- ESCO revenue includes only contractor fees and maintenance profits.
- Depreciation calculated linearly over investment cost.
- Residual value considered at the end of the analysis period.
- Energy company subsidiary responsible for operation and maintenance.
- O&M costs on the energy company's side are 90% of O&M fees.
- O&M costs saved due to modernization: 25%
- Identical external conditions: VAT = 27%, inflation = 3.00%, required return = 5.83%, corporate tax = 9%, analysis period = 15 years.
- Identical energy, transmission, and distribution tariffs for old and new systems.
- Tariffs increase according to inflation.

Table 1: Description of the current system

System elements	Quantity (Nr.)	Electrical energy (W)	Consumption (Wh)
Na lamp 400W	9	400	432
Na lamp 70W	44	70	79
Metal halide lamp 70W	10	70	70
Compact fluorescent lamp 36W	352	36	45
SUM	**415**	**20 052**	**23 904**

Table 2: Description of the planned system

System elements	Quantity (Nr.)	Electrical energy (W)	Consumption (Wh)	Rated lifetime
Na lamp 400W	9	200	210	60 000
Na lamp 70W	44	55	59	60 000
Metal halide lamp 70W	10	55	59	60 000
Compact fluorescent lamp 36W	352	30	33	60 000
SUM	**415**	**15 330**	**16 692**	

3 Results and Discussion

3.1 Financial Results for the business options

The financial analysis used traditional indicators such as NPV, IRR, payback period, and ROI. Without ESCO financing (see Table 3), modernizing public lighting would require upfront costs for the municipality. The first year shows a net cost of EUR 48,918.78, based on an EUR 52,675 investment, only partly offset by savings from reduced electricity and 25% lower O&M costs due to improved energy efficiency. Despite operational savings, all key financial indicators remain negative, and the project does not cover initial costs over 15 years. For the energy company, the NPV of revenues from operational and maintenance cash flows, excluding electricity sales, is about EUR 2253.513.

Table 3. Financial results for the partners under Option 1

Indicator	Municipality	Energy Company
Total discounted savings/revenue (EUR)	-1 824.295	2253.513
Yearly discounted savings (EUR/y)	-121.62	150.24
Discounted internal rate of return	-0.53%	Not interpretable
Discounted payback period	Not interpretable	0
ROI	-3.00%	Not interpretable

Under the ESCO model, the municipality collaborates with an energy company to modernize public lighting by replacing existing lighting with LED luminaires. The energy company finances the upgrade, and the municipality repays over an eight-year period, starting from the second year, with an annual fee of 1% of the investment. The project reaches financial breakeven in year 12, with total discounted savings of EUR 6,354.21. From the energy company's view, the investment does not break even within 15 years, resulting in a net present value of EUR -6,150.27 and a discounted internal rate of return of -3.92%.

Table 3. Financial results for the partners under Option 2

Indicator	Municipality	Energy Company
Total discounted savings (EUR)	6354.205	-6150.27
Yearly discounted savings (EUR/y)	423.61	410.02
Discounted internal rate of return	5.66%	-3.92%
Discounted payback period	12	Not interpretable
ROI	58.00%	Not interpretable

In sum, the financial analysis indicates that, under the baseline assumptions, the LED modernization of the public lighting system is not economically viable when financed solely by the municipality. Under the ESCO model, the project becomes financially attractive to the municipality but fails to generate a sufficient return for the energy company. However, these results are highly sensitive to a range of external factors that are uncertain at the start of the project or may evolve over time, as well as to contractual parameters that may be subject to negotiation between the parties. Consequently, conclusions drawn from the baseline options alone are insufficient to provide robust and well-founded guidance for decision-making.

3.2 Financial results for the business options

To illustrate how traditional and ESCO models respond to changes in assumptions on exogenous variables or contractual elements, one and two-way sensitivity analyses were conducted. The existing literature stresses that public lighting retrofit projects are sensitive to energy charges (Beccali et al., 2018; Campisi et al., 2017), subsidies (World Bank, 2015), contractor fees, and contract length (Guo & Zhang, 2020; Tupikina & Rozhkova, 2020; Martiniello et al., 2020). Energy charge changes directly impact savings: higher charges boost benefits, while steady or falling prices weaken financial viability. Grant support reduces initial costs and improves financial indicators. One-way analyses on energy charges and grants were performed for Options 1 and 2. As shown in Figure 2, under Option 1, municipal savings increase if energy charges fall below EUR 0.03827/kWh. With grants, savings break even at a 3.46% grant rate.

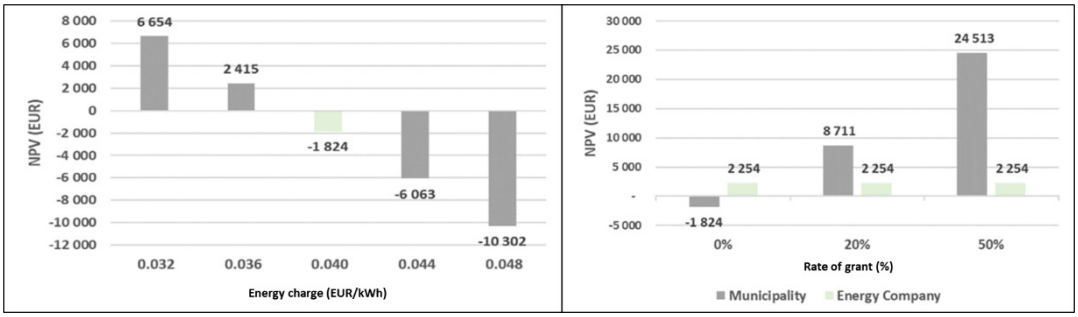


Figure 2. Sensitivity analysis for Option 1 - Energy charge (EUR/kWh) & Availability of public grants (Rate of grant in the investment cost, %)

In the case of Option 2 (see Figure 3), the municipality breaks even over the evaluation period at an energy charge of EUR 0.046/kWh. Beyond this level, the investment is no longer financially beneficial for the municipality. This suggests that higher energy charges alone are sufficient to eliminate municipal financial gains. The availability of public grant improves the economic results for both parties, but not to the same extent. Even at a high subsidy rate of 50%, the energy company is still unable to fully recover its investment over the evaluation period, although its net present value improves to EUR -1,372. By contrast, the municipality benefits more strongly from grant funding. Under these conditions, the investment is effectively recovered in the year of implementation, resulting in total discounted cost savings of EUR 28,602.

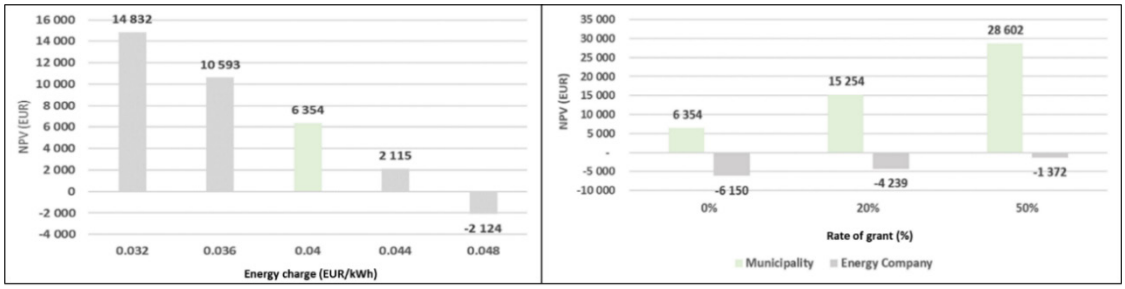


Figure 3. Sensitivity analysis for Option 2 - Energy charge (EUR/kWh) & Availability of public grants (Rate of grant in the investment cost, %)

The municipality's ESCO financial results are sensitive to contractual parameters. Small changes in fees can significantly impact net present savings: a 1% fee increase without grants can make the NPV negative. Conversely, this boosts the energy company's NPV to EUR -4,488.3. The project becomes profitable for the energy company once the contractor fee exceeds 4.63%, while municipality savings turn negative above 3.77% (Figure 4). Contract duration and initial repayment rate are key, affecting how returns are shared. Longer contracts ease annual pressure and help recover costs, while higher initial repayments improve the energy company's liquidity but increase early costs for the municipality. Longer repayment periods with higher first-year payments can increase the municipality's discounted savings but reduce the energy company's return.

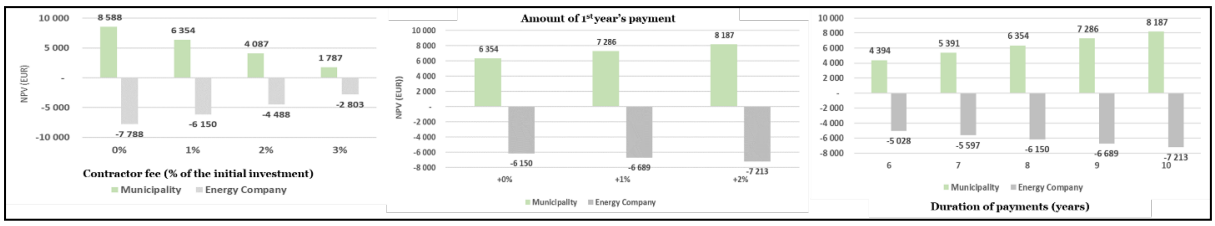


Figure 4. Sensitivity analysis for Option 2 - Contractor fee, Amount of 1st year's payment & Duration of payments

The two-way sensitivity analysis (see Figure 5) examines how changes in public grants and contractor fees affect the project's net present value, helping determine when investment is attractive to municipalities without harming the ESCO's finances. Results show that with only contractual fee and O&M income, and no grants, no contractor fee yields positive returns for both parties, revealing a conflict of interest in the ESCO model without public support. However, if subsidies exceed 30%, municipal savings stay positive across all fees, showing grants can turn an unviable project into a financially sustainable one, enabling mutually acceptable outcomes.

Municipality												
		Availability of public grants (%)										
		0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
Contractor fee (%)	0%	8 588	12 815	17 041	21 267	25 493	29 719	33 946	38 172	42 398	46 624	50 851
	1%	6 354	10 804	15 254	19 703	24 153	28 602	33 052	37 502	41 951	46 401	50 851
	2%	4 087	8 763	13 440	18 116	22 792	27 469	32 145	36 822	41 498	46 174	50 851
	3%	1 787	6 693	11 600	16 506	21 412	26 319	31 225	36 132	41 038	45 944	50 851
	4%	-546	4 594	9 734	14 873	20 013	25 153	30 292	35 432	40 571	45 711	50 851
	5%	-6 094	-657	4 779	10 216	15 653	21 090	26 526	31 963	37 400	42 837	48 273
	6%	-8 511	-2 833	2 846	8 524	14 203	19 881	25 559	31 238	36 916	42 595	48 273
	7%	-10 959	-5 036	887	6 810	12 734	18 657	24 580	30 504	36 427	42 350	48 273
	8%	-13 438	-7 267	-1 095	5 076	11 247	17 418	23 589	29 760	35 931	42 102	48 273
	9%	-8 445	-9 524	-3 102	3 320	9 742	16 164	22 586	29 008	35 430	41 851	48 273
	10%	-18 482	-11 807	-5 131	1 544	8 220	14 895	21 571	28 247	34 922	41 598	48 273

Net Present Value at the energy company's side (EUR)												
		Availability of public grants (%)										
		0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
Contractor fee (%)	0%	-7 788	-6 669	-5 550	-4 430	-3 311	-2 191	-1 072	-206	614	1 434	2 254
	1%	-6 150	-5 195	-4 239	-3 284	-2 328	-1 372	-417	322	966	1 610	2 254
	2%	-4 488	-3 699	-2 910	-2 120	-1 331	-541	248	857	1 323	1 788	2 254
	3%	-2 803	-2 182	-1 561	-940	-320	301	922	1 401	1 685	1 969	2 254
	4%	-1 094	-644	-194	256	706	1 156	1 606	1 951	2 052	2 153	2 254
	5%	637	914	1 191	1 467	1 744	2 021	2 298	2 510	2 424	2 333	2 254
	6%	2 390	2 492	2 593	2 695	2 796	2 898	2 999	3 076	2 802	2 528	2 254
	7%	4 166	4 090	4 014	3 938	3 862	3 786	4 938	3 633	3 184	2 719	2 254
	8%	5 962	5 707	5 451	5 195	4 940	4 684	3 710	4 172	3 570	2 912	2 254
	9%	7 780	7 342	6 905	6 468	6 030	5 593	5 155	4 718	3 962	3 108	2 254
	10%	9 618	8 997	8 376	7 754	7 133	6 512	5 766	5 269	4 357	3 305	2 254

Figure 5. Sensitivity analysis for Option 2 - Public grants & Contractual fee

4 Discussion

In the relevant literature, only a limited number of studies explicitly examine municipal and ESCO cash flows in parallel or compare direct public financing with ESCO-based delivery using a unified discounted cash flow framework. According to prior studies, municipalities can realize considerable energy savings through street lighting modernization despite failing to recoup investment costs (Tähkämö et al., 2012; Beccali et al., 2018). It can also be stated that these economically desirable outcomes are highly context-specific and depend on local tariff structures, engineering and maintenance assumptions, lighting standards, capital costs, and financing conditions (Jiang et al., 2015; Campisi et al., 2018). The results of the present study are fully consistent with these statements, confirming that under self-financing, financial savings often materialize too slowly to offset upfront costs, particularly in the absence of subsidies or favorable tariffs (Polzin et al., 2016a; Novikova et al. 2017). At the same time, the relevant literature treats ESCO-based models as a means of improving public street lighting due to their predictable savings, moderate technical complexity, and relatively reliable baseline consumption (Šeba & Mušec, 2018). Nevertheless, both prior findings and our results emphasize that ESCO financing does not eliminate costs but redistributes them over time through contractor fees, contract duration, and risk allocation (Tupikina & Rozhkova, 2018; Martiniello et al., 2020). Consistent with earlier empirical and modeling research, our study shows that low contractor fees may preserve positive municipal savings while undermining ESCO viability, whereas small increases in fees or contract length can rapidly shift returns toward service providers (Guo et al., 2019; Guo & Zhang, 2020). Overall, these findings reinforce contract design and savings sharing as the central determinants of balanced and sustainable ESCO business models.

5 Conclusions

Although the technical effectiveness of LED street-lighting retrofits is well established, their financial performance becomes less certain once financing arrangements and contract design are considered. This study uses a two-sided discounted cash-flow approach to compare direct public financing with ESCO-based implementation. The results show that viability depends not only on energy savings, but also on tariffs, contract duration, fees, and subsidies. Even energy-efficient LED projects may be financially unattractive under direct financing, while ESCO models can relieve municipal budget pressure, often at the expense of provider profitability. In the examined cases, balanced outcomes for both parties occur only when public investment subsidies exceed 30 percent. Generalization of our key findings is not possible because the analysis is based on a hypothetical Hungarian case with a specific technical, economic, and institutional context. Although exploratory and context-specific, the analysis helps clarify when ESCO financing can offer a financially sustainable path for public lighting modernization.

6 Ethics Declaration

Ethical approval was not required for the research reported in this paper.

7 AI Declaration

During the preparation of this work, the authors used Grammarly and ChatGPT to improve the study's readability and language. After using this tool/service, the authors reviewed and edited the content as needed.

References

Aasen, M., Westskog, H. and Korneliussen, K. (2016) “Energy performance contracts in the municipal sector in Norway: overcoming barriers to energy savings?”, *Energy Efficiency*, Vol. 9, No. 1, pp. 171-185.

Balázs, L., Braun, F., and Lengyel, J. (2023) “Energy saving potential of traffic-regulated street lighting”, *Sustainability*, Vol. 15, No. 8, 6750.

Beccali, M., Bonomolo, M., Leccese, F., Lista, D., and Salvadori, G. (2018) “On the impact of safety requirements, energy prices and investment costs in street lighting refurbishment design”, *Energy*, Vol. 165, pp. 739-759.

Bertoldi, P., Boza-Kiss, B., Panev, S., and Labanca, N. (2014) “The European ESCO Market Report 2013”, Science and Policy Report, Final report.

Bertoldi, P., & Rezessy, S. (2005) Energy service companies in Europe. Status report, pp. 1-72.

Bobbino, S., Galván, H., and González-Eguino, M. (2014) *Budget-neutral financing to unlock energy savings potential: An analysis of the ESCO model in Barcelona*. In: Green energy and efficiency: An economic perspective (pp. 183-199). Springer International Publishing, Cham.

Campisi, D., Gitto, S., and Morea, D. (2018) “Economic feasibility of energy efficiency improvements in street lighting systems in Rome”, *Journal of Cleaner Production*, Vol. 175, pp. 190-198.

Campisi, D., Gitto, S., and Morea, D. (2017) "Light emitting diodes technology in public light system of the municipality of Rome: An economic and financial analysis", *International Journal of Energy Economics and Policy*, Vol. 7, No. 1, pp. 200-208.

Carli, R., Dotoli, M. and Pellegrino, R. (2018) "A decision-making tool for energy efficiency optimization of street lighting", *Computers & Operations Research*, Vol. 96, pp. 223-235.

Doulos, L. T., Sioutis, I., Kontaxis, P., Zissis, G., and Faidas, K. (2019) "A decision support system for assessment of street lighting tenders based on energy performance indicators and environmental criteria: Overview, methodology and case study", *Sustainable Cities and Society*, Vol. 51, 101759.

European Commission (2025): *Assessment of Hungary's updated NECP*, [online], https://commission.europa.eu/publications/hungary-final-updated-necp-2021-2030-submitted-2024_en.

Gellén, M., and Horváth, A. (2026) "Digital Municipalities in Hungary: Examples of Participation and Engagement in the Orbit of the European Union's Digital Strategies", *Public Finance and Management*, 15239721261418721.

Gordic, D., Vukasinovic, V., Kovacevic, Z., Josijevic, M. and Zivkovic, D. (2021) "Assessing the techno-economic effects of replacing energy-inefficient street lighting with LED corn bulbs", *Energies*, Vol. 14, No. 13, 3755.

Guo, K., and Zhang, L. (2020) "Guarantee optimization in energy performance contracting with real option analysis", *Journal of Cleaner Production*, Vol. 258, 120908.

Guo, K., Zhang, L., and Wang, T. (2019) "Optimal scheme in energy performance contracting under uncertainty: a real option perspective", *Journal of Cleaner Production*, Vol. 231, pp. 240-253.

Hermoso-Orzáez, M. J., Lozano-Miralles, J. A., Lopez-Garcia, R., and Brito, P. (2019) "Environmental criteria for assessing the competitiveness of public tenders with the replacement of large-scale LEDs in the outdoor lighting of cities as a key element for sustainable development: Case study applied with PROMETHEE methodology", *Sustainability*, Vol. 11, No. 21, 5982.

Jensen, J. O., Nielsen, S. B., and Hansen, J. R. (2013) "Greening public buildings: ESCO-contracting in Danish municipalities", *Energies*, Vol. 6, No. 5, pp. 2407-2427.

Jiang, Y., Li, S., Guan, B., and Zhao, G. (2015) "Cost effectiveness of new roadway lighting systems", *Journal of traffic and transportation engineering*, Vol. 2, No. 3, pp. 158-166.

Juričić, D., Vašiček, D. and Drezgić, S. (2020) "Multiple criteria decision analysis of public investment options: application to streetlighting renewal projects", *Economic research-Ekonomska istraživanja*, Vol. 33, No. 1, 3288-0.

Kurniawan, I. W. and Kurniawan, R. (2022) "Fostering energy efficiency through street lighting system improvement: a case study of city-level in Indonesia", *Indonesian Journal of Energy*, Vol. 5, No. 2, pp. 96-104.

Kurowska-Pysz, J. and Kunikowski, G. (2021) "The ESCO formula as support for public and commercial energy projects in Poland", *Energies*, 14.23: 8098.

Lockwood, R., and Selwyn, T. (2011) A review of local authority road lighting initiatives aimed at reducing costs, carbon emissions and light pollution. Research Report, [online], www.templegroup.co.uk

Martiniello, L., Morea, D., Paolone, F., and Tiscini, R. (2020) "Energy performance contracting and public-private partnership: How to share risks and balance benefits", *Energies*, Vol. 13, No. 14, 3625.

Marques, R. C. and Geddes, R. R. (2019) "The use of PPP arrangements in street lighting: A win-win option?", *Annals of Public and Cooperative Economics*, Vol. 90, No. 2, pp. 311-327.

Mercado, J., Martin, B., and Marco, F. (2018) How to procure dynamic lighting projects: A practical guide, Interreg Research Report, [online], https://www.ikem.de/wp-content/uploads/2021/07/D.T4.2.3-Strategy-to-facilitate-the-integration-of-dynamic-lighting-from-a-legal-perspective.pdf?utm

Meričková, B. M., Muzhová, N. J., and Holúbek, M. (2020) A Comparison of PPPs and Traditional Provision of Public Serviicess at the Municipal Level, XXIII. mezinárodní kolokvium o regionálních vědách, 207. , [online], https://www.researchgate.net/profile/Martin-Lusticky-3/publication/342388776_The_dynamics_of_tourism_destination_competitiveness/links/60b1f89c45851557baa6f228/The-dynamics-of-tourism-destination-competitiveness.pdf#page=207.

Novikova, A., Stamo, I., Stelmakh, K., and Hessling, M. (2017) Guideline on finding a suitable financing model for public lighting investment. Interreg Central Europe, Dynamic Light Project, [online], https://www.porsennaops.cz/uploads/media/default/0001/01/b98d834f21d4c7848bc431d5776105e712880b96.pdf.

OECD (2023) Promoting Nature-based Solutions in Municipalities in Hungary, [online], https://www.oecd.org/content/dam/oecd/en/publications/reports/2023/12

Patra, O., Ilyas, F. F. and Suryana, H. (2025) "Feasibility Study on The Implementation of Public-Private Partnership: Energy Conservation Based on Value for Money and Project Risk. Return: Study of Management", *Economic and Business*, Vol. 4, No. 9, pp. 748-755.

Pintér, G., Baranyai, N. H., Wiliams, A., and Zsiborács, H. (2018) "Study of photovoltaics and LED energy efficiency: Case study in Hungary", *Energies*, Vol. 11, No. 4, 790.

Plepys, A., and Richter, J. L. (2016) "Public procurement barriers in promoting market uptake of innovative LED lighting. In 2016 Electronics Goes Green 2016+(EGG)", (pp. 1-8). IEEE.

Polzin, F., von Flotow, P., and Nolden, C. (2016a) "Modes of governance for municipal energy efficiency services–The case of LED street lighting in Germany", *Journal of Cleaner Production*, Vol. 139, pp. 133-145.

Polzin, F., von Flotow, P. and Nolden, C. (2016b) "What encourages local authorities to engage with energy performance contracting for retrofitting? Evidence from German municipalities", *Energy Policy*, Vol. 94, pp. 317-330.

Polzin, F., Nolden, C., and von Flotow, P. (2018) “Drivers and barriers for municipal retrofitting activities–Evidence from a large-scale survey of German local authorities”, *Renewable and Sustainable Energy Reviews*, Vol. 88, pp. 99-108.

Šeba, M. G., and Mušec, M. (2018) *Making Street Lighting ESCO Projects Work in Practice*. In Economy, Finance and Business in Southeastern and Central Europe: Proceedings of the 8th International Conference on the Economies of the Balkan and Eastern European Countries in the Changing World (EBEEC) in Split, Croatia, 2016 (pp. 257-271). Cham: Springer International Publishing.

Tähkämö, L., Ylinen, A., Puolakka, M., and Halonen, L. (2012) “Life cycle cost analysis of three renewed street lighting installations in Finland”, *The International Journal of Life Cycle Assessment*, Vol. 17, No. 2, pp. 154-164.

Tsankov, P., Yovchev, M., and Ibrishimov, H. (2025) “Energy-efficient light-emitting diode retrofit and advanced control in municipal street lighting: A case study from Bulgaria”, *Asian Journal of Water*, *Environment and Pollution*, Vol. 22, No. 6, pp. 180-197.

Tupikina, A. A., and Rozhkova, M. V. (2018) *Economic evaluation of energy service contract implementation from the view of its participants*. In 2018 XIV International Scientific-Technical Conference on Actual Problems of Electronics Instrument Engineering (APEIE) (pp. 459-463). IEEE.

Valentiny, P. (2008) “Energy services at local and national level in the transition period in Hungary (No. MT-DP-2008/4)”, IEHAS Discussion Papers.

Vasvári, T., and Longauer, D. (2024) “Against the tide: how changes in political alignment affect grant allocation to municipalities in Hungary”, *Public Choice*, Vol. 198, No. 3, pp. 467-492.

Vidovszky-Németh, A., and Schanda, J. (2011) “Investigation of Street Lighting with White LEDs in Hungary”, *Light and Engineering*, Vol. 19, No. 2, pp. 5-13.

Wacinkiewicz, D. and Słotwiński, S. (2023) “The Statutory Model of Energy Performance Contracting as a Means of Improving Energy Efficiency in Public Sector Units as Seen in the Example of Polish Legal Policies”, *Energies*, Vol. 16, No. 13, 5060.

Wardhana, Y. M. A. (2024) “The Provision of Public Street Lighting Based on Risk Mitigation for Energy Efficiency and Environmental Protection”, *Prisma Sains: Jurnal Pengkajian Ilmu dan Pembelajaran Matematika dan IPA IKIP Mataram*, Vol. 12, No. 1, pp. 148-159.

World Bank (2015) India - Energy-efficient Street lighting: implementation and financing solutions, Report, [online], https://documents1.worldbank.org/curated/en/444081468282557427/pdf/India-EE-Street-Lighting-Implementation-and-Financing-P149482-June-27-2015.pdf.

Path Development and the Twin Transition in Europe

Vincenzo Maione[1], Cristina Ponsiglione[2], Simonetta Primario[3]
[1] LIUC – Cattaneo University, Castellanza, VA, Italy
[2] University of Naples Federico II, Naples, Italy
[3] Pegaso Telematic University, Italy

Abstract

A central challenge in regional innovation and sustainability transitions research is to explain why the twin innovation transition — understood as the joint advancement of green and digital technological specialisation — emerges only in some regions and only at specific moments in time. While Evolutionary Economic Geography has offered powerful concepts to analyse regional path development, much of this literature has focused on the emergence of new growth trajectories, paying less attention to situations in which path-development potential stabilises, remains latent, episodic, or insufficiently converted into transformative innovation outcomes.

This paper addresses this gap by developing a configurational panel framework that links regional path development mechanisms to twin innovation outcomes across European NUTS-2 regions over the period 2007–2019. Drawing on regional path development theory and sustainability transitions research, we argue that twin innovation is not the result of a single enabling factor, but of place-specific combinations of structural assets, innovation capabilities, external linkages, policy support, and institutional conditions. At the same time, the absence of twin innovation should not be interpreted simply as a lack of capabilities, but may reflect incomplete conversion processes, weak cross-domain coupling, or path-development dynamics that remain temporally unstable.

Methodologically, the paper applies panel fuzzy-set Qualitative Comparative Analysis to distinguish between cross-sectional variation across regions and longitudinal variation within regions. This allows us to assess not only which configurations are sufficient for twin innovation, but also whether they display temporal stability or instead appear as episodic alignments. The findings show that regional innovation performance constitutes an important baseline condition across the identified configurations, but is not sufficient on its own to generate twin innovation outcomes. Moreover, the mechanisms most commonly associated with cumulative regional development, such as extension and branching, do not emerge as sufficient pathways to twin innovation. Instead, the results point to a limited number of highly place-specific configurations, often concentrated in Northern and Western European regions, whose empirical relevance is frequently episodic rather than stable over time.

The paper contributes to the literature in two ways. Conceptually, it refines the link between regional path development and sustainability transitions by distinguishing between path-development potential, path activation, and path realisation. Empirically, it provides systematic panel evidence on the spatial selectivity and temporal fragility of twin innovation trajectories in Europe, with implications for place-based innovation policies aimed at supporting not only regional capabilities, but also their effective conversion into green-digital transformation.

Keywords: Twin Transition, Regional path development, Multi-method Research, Panel fsQCA, Process Tracing.

Paper type: Academic Research Paper

1 Introduction

Regional economies across Europe have undergone substantial transformation over the past two decades, driven by intensifying pressures towards green and digital technological change (Bachtrögler-Unger et al., 2023). Among the multiple dimensions of this transformation, the joint advancement of green and digital technological specialisation within regions — what we term twin innovation — plays a particularly strategic role. It is through the simultaneous generation and adoption of knowledge in both domains that regions build the capabilities required to navigate long-term structural change (Mäkitie et al., 2023). Yet this process remains deeply uneven across space and over time. A growing body of research has begun to examine regional innovation trajectories, highlighting how structural heterogeneity shapes both the capacity and the mode of regional response to transformative pressures (Fazio et al., 2024; Montresor & Vezzani, 2023). While this heterogeneity has been examined in its spatial dimension (R. Boschma et al., 2017), its temporal dynamics — including whether underlying configurations persist, transform, or emerge only episodically — have largely remained a black-box in transitions research (Strambach & Pflitsch, 2020).

Evolutionary Economic Geography provides a powerful framework for analysing how regional economies transform their innovation base over time (R. A. Boschma & Frenken, 2006; Martin & Sunley, 2006). Its typology of path development mechanisms — creation, importation, upgrading, diversification, branching, and extension — offers a systematic basis for understanding regional structural change (Grillitsch et al., 2018; Hassink et al.,

2019; Isaksen, 2020). However, this literature has predominantly focused on the emergence of new growth trajectories, paying limited attention to situations in which path-development potential remains latent, episodic, or insufficiently converted into transformative innovation outcomes. More fundamentally, the link between regional path development mechanisms and twin innovation trajectories has not been systematically established — neither for configurations that succeed, nor for those that remain stable yet fail to generate green-digital co-specialisation over time.

This paper addresses these gaps by developing a configurational panel framework linking regional path development mechanisms to twin innovation outcomes across European NUTS-2 regions over 2007–2019. By combining insights from sustainability transitions theory and Evolutionary Economic Geography (Fuenfschilling & Truffer, 2014; Köhler et al., 2019; Loorbach et al., 2017), and applying panel fuzzy-set Qualitative Comparative Analysis, the analysis disentangles cross-sectional structural diversity from temporal dynamics, highlighting the extent to which regional pathways reflect durable structures or transient alignments of conditions.

The remainder of the paper is structured as follows. Section 2 reviews the theoretical background. Section 3 presents the methodology. Section 4 presents the analytical model. Section 5 reports the empirical findings. Section 6 discusses their theoretical and policy implications and concludes the paper.

2 Background

Understanding why twin innovation trajectories unfold so unevenly across regions requires unpacking the place-specific configurations that shape regional innovation systems and their capacity for transformation over time. Regions can be conceptualised as territorially embedded systems of firms, knowledge organisations, and institutions whose interactions shape processes of knowledge creation, diffusion, and use (Asheim et al., 2011). These systems are embedded in multi-scalar dynamics that connect local developments to broader socio-technical transformations (Coenen et al., 2012; Truffer & Coenen, 2012), and are shaped by institutional environments that co-evolve with technological specialisation and industrial structures. As a result, regional transitions tend to follow differentiated and context-specific trajectories, reflecting place-dependent configurations of capabilities and institutional arrangements (Affolderbach & Schulz, 2016; Gibbs & O'Neill, 2017). A growing body of work at the intersection of sustainability transitions research and Evolutionary Economic Geography has highlighted precisely this multiplicity — emphasising how the pace, direction, and stability of regional transformation are shaped by place-specific configurations rather than by universal mechanisms (R. Boschma et al., 2017; Hodson et al., 2017).

The EEG literature provides a well-established framework for analysing how regional economies evolve through processes of path development, emphasising the cumulative and path-dependent nature of regional transformation shaped by the interaction between existing capabilities, institutional structures, and external linkages (R. Boschma & Martin, 2010). The literature identifies a range of mechanisms through which regional industrial and innovation structures evolve over time — path creation, importation, upgrading, diversification, branching, and extension (Grillitsch et al., 2018; Hassink et al., 2019) — as synthesised in Table 1.

Table 1. Types and mechanisms of regional industrial path development

Type of path development	Definition
Path creation	Emergence of entirely new industries, driven by novel technologies, institutional change, and entrepreneurial agency, often requiring the transformation of the local knowledge and innovation system.
Path branching	Emergence of new industries related to existing regional knowledge, based on the redeployment of existing assets by incumbent firms or spin-offs.
Path diversification	Emergence of new industries based on unrelated combinations of knowledge, involving a greater degree of novelty and cognitive distance.
Path importation	Establishment of industries new to the region but not new to the world, driven by exogenous actors and potentially embedded through local linkages (Karlsson, 2008).
Path upgrading	Transformation within existing paths via (i) functional upgrading in global value chains, (ii) technological or organisational renewal, or (iii) niche development based on symbolic knowledge.
Path extension	Continuity of existing industrial path through incremental innovation along established technological trajectories

Source: own elaboration

Yet this classification, while analytically useful, tends to conceptualise path development mechanisms as distinct and relatively stable processes. To assess the empirical grounding of this limitation, we conducted a

structured review of studies explicitly applying path development typologies across European regional contexts — presented in detail in Section 3. The review confirms that existing studies have largely adopted descriptive approaches focused on a limited number of case studies, offering only partial insight into how multiple mechanisms interact, overlap, and reconfigure over time (Blažek et al., 2020). This empirical orientation contrasts with the sustainability transitions literature, which conceptualises transformation as the outcome of co-evolving technological, institutional, and social dynamics unfolding through region-specific configurations of industrial structures and actor constellations (Geels, 2012; Hansen & Coenen, 2015; Markard et al., 2012; Truffer & Coenen, 2012). What remains insufficiently understood is how multiple path development mechanisms combine and evolve over time — and whether the resulting configurations are durable structures or transient alignments that dissolve before generating transformative outcomes.

This gap is particularly salient in the context of twin innovation, which requires the simultaneous development and coupling of green and digital knowledge domains within the same regional system over time. While both domains have been widely studied in isolation, the link between regional path development mechanisms and twin innovation trajectories has not been systematically established — neither for configurations that succeed in generating green-digital co-specialisation, nor for those that remain stable yet fail to convert path-development potential into twin innovation outcomes. Against this background, the paper addresses the following research question: How do configurations of regional path development mechanisms evolve over time across European regions, and what conditions determine whether these temporal patterns support, constrain, or prevent the emergence of twin innovation trajectories?

3 Methodology

Investigating how configurations of regional path development mechanisms evolve over time and under which conditions they support, constrain, or prevent twin innovation trajectories requires a methodological approach capable of addressing two forms of complexity simultaneously: the geographical heterogeneity of regional systems and the temporal dynamics of their transformation. Neither regression-based models nor structural equation modelling are well suited to this purpose, as they estimate net effects of individual conditions and are ill-equipped to capture conjunctural causation, equifinality, and causal asymmetry (Ragin, 2000, 2008). The analysis therefore adopts a two-step design that combines a structured literature review with a panel fuzzy-set Qualitative Comparative Analysis (fsQCA).

The first step consists of a structured review of empirical studies explicitly applying path development typologies across regional contexts. Following a multi-stage protocol, we queried Scopus using a combination of relevant keywords — including "regional innovation", "path development", "branching", "diversification", "upgrading", "importation", and "creation" — restricted to peer-reviewed articles in English. Abstracts were screened for references to European countries and NUTS-2 regions, and studies from unrelated disciplines were excluded. Additional papers were included based on the research team's expert judgement. The resulting cases were systematically classified according to their dominant path development trajectory by applying the taxonomy drawn from Grillitsch et al. (2018) to the mechanisms explicitly described by the original authors. This step provided the empirical foundation for the second step.

The second step adopts a configurational approach grounded in a conceptualisation of regional innovation systems as complex adaptive systems — dynamic structures characterised by non-linearity and emergent properties arising from interactions among heterogeneous agents (Cooke, 2013). FsQCA builds on set-theoretical principles and Boolean algebra, and draws on an understanding of causal complexity structured around conjunctural causation, equifinality, and causal asymmetry (Rihoux & Ragin, 2009). Unlike regression approaches, fsQCA identifies how multiple conditions combine in shaping an outcome rather than estimating their net individual effects — a distinction that is particularly relevant for economic geography, where conditions often matter only in conjunction with others (Woodside, 2013). Causal asymmetry is equally important here: the configurations associated with the presence of twin innovation outcomes are not simply the mirror image of those associated with their absence (Fiss et al., 2013).

To extend this configurational logic to the temporal dimension, the analysis adopts a panel fsQCA approach, which extends set-theoretic reasoning to longitudinal data and makes it possible to distinguish cross-sectional variation across regions from longitudinal variation within regions (Garcia-Castro & Ariño, 2016). This makes it possible to assess not only which configurations are sufficient for twin innovation, but also whether they display temporal stability or appear as episodic alignments. Following Bhattacharya (2023), panel-data QCA modelling proceeds through case knowledge, model specification, set calibration, directional expectations, and tests of necessity and sufficiency. Pooled, between, and within consistency and coverage scores are computed to distinguish cross-regional from within-region temporal variation.

4 The model

This section presents the analytical model used to investigate how combinations of regional path development mechanisms shape twin innovation outcomes across European NUTS-2 regions over 2007–2019.

The selection of causal conditions draws on both theoretical insights from EEG and the review presented in Section 3. Table 2 presents the full set of causal conditions, their operational definitions, and data sources.

Table 2. Causal Conditions

Label	Dimension	Index	Source
UV	Unrelated Variety	Unrelated Variety	Eurostat
EKIA	Structural complexity	Employment in Knowledge-intensive activities	Regional Innovation scoreboard
FDI	FDI Intensity	FDI inward stock % GDP	OECD
ICOLL	Collaborative R&D activity	Innovative SMEs collaborating with others	Regional innovation Scoreboard
PS	Policy Support	Historic EU payments annual timeseries - modelled	Cohesion Data
RII	Innovation Performance	Regional Innovation Index	Regional Innovation Scoreboard
NEWB	Entrepreneurial Activity	Business entries	OECD Data Explorer

The outcome is constructed using Revealed Technological Advantage (RTA), a proxy for regional technological specialisation derived from patent data (Balland et al., 2019). RTA captures the extent to which a region is oriented towards a specific technological domain relative to the overall European patenting structure. Yearly RTAs were calculated using OECD REGPAT data. Green technologies refer to the OECD "environment-related technologies" domain, while digital technologies include patents classified under "ICT" and "artificial intelligence-related technologies". RTA is calculated as follows:

$$RTA_{r,d} = \frac{P_{r,d}}{P_r} \Big/ \frac{P_d}{P_T}$$

where:

- $P_{r,d}$ is the number of patent applications in domain d filed in region r,
- P_r is the total number of patent applications filed in region r,
- P_d is the total number of patent applications in domain d across all regions,
- P_T is the total number of patent applications in all domains and regions.

The twin innovation outcome is constructed at the region-year level as the minimum between green and digital RTA:

$$RTA_Twin_{r,t} = \min\left(RTA_Digital_{r,t}, RTA_Green_{r,t}\right)$$

This reflects the directional and conjunctural nature of twin innovation: a regional transition path can be considered twin-oriented only when green and digital specialisation are jointly present. The resulting values are calibrated into fuzzy-set membership scores, while temporal stabilisation or fragility is assessed through the panel fsQCA analysis.

To ensure the reliability of RTA values, we applied a two-level filtering strategy. At the country level, only countries with at least 1,000 patent applications over the reference period were retained, following Cantwell (2010) and Vertova (2001). At the regional level, we excluded NUTS-2 regions with fewer than 50 patent applications per year and required at least 10 patents in either the green or digital domain, adapting the criteria proposed by Pinheiro et al. (2022) to the twin innovation context. These filters reduce small-number distortions while preserving regions specialised in at least one transition domain. The final dataset includes 82 NUTS-2 regions across 12 European countries for 2007–2019.

The regional nomenclature follows the classification used in the Regional Innovation Scoreboard.

5 Results

The necessity analysis shows that no individual condition reaches the conventional consistency threshold of 0.90. Regional innovation performance has the highest consistency value (0.815), but remains below the threshold, indicating that it is relevant but not necessary for twin innovation. This supports the need for a configurational analysis. Summary results are reported in Table 4.

Table 3. Necessity analysis for the twin transition outcome.

Causal Condition	Consistency	Coverage
EKIA	0,733	0,529
ICOLL	0,708	0,526
NEWB	0,635	0,486
FDI	0,695	0,503
UV	0,757	0,514
PS	0,625	0,491
RII	0,815	0,498

The sufficiency analysis identifies six configurations associated with the twin innovation outcome, with overall coverage of 0.91 and consistency of 0.75. Table 3 reports the configurations, fit parameters, covered cases, and path-development interpretation. Regional innovation performance (RII) appears as a core condition across all six configurations, confirming its role as a baseline structural requirement — a necessary floor below which no configuration is sufficient, but which alone does not generate twin innovation outcomes.

The combination of moderate raw coverage and very low unique coverage scores — ranging from 0.002 to 0.029 — indicates that the identified configurations cover largely overlapping sets of cases, each explaining only a marginal share of the outcome independently. More importantly, the outcome membership scores of covered cases reveal that the large majority display weak twin innovation membership, typically between 0.10 and 0.40. These are better understood as deviant consistency cases in the sense of Oana and Schneider (2018): regions whose structural conditions align with a sufficient configuration, yet fail to convert that alignment into realised twin innovation outcomes. Cases with high outcome membership are rare, geographically concentrated in a small number of Northern and Western European regions, and tend to appear only in isolated time windows. This pattern constitutes the empirical signature of what this paper terms episodic alignment — configurations that are formally sufficient yet dissolve before stabilising into durable twin innovation trajectories across institutional fields and over time (Strambach & Pflitsch, 2020).

This prevalence of deviant consistency cases carries broader theoretical meaning. It suggests that the absence of twin innovation should not be treated as a residual category, but as an empirical signal of path-development dynamics that remain weakly converted, structurally blocked, or insufficiently coupled across domains. In this sense, the findings speak directly to the debate on the dark side of regional path development (Blažek et al., 2020): while the analysis does not directly measure trajectories of decline, it shows that identifiable path-development mechanisms may still produce incomplete or non-transformative transition trajectories when they fail to generate the simultaneous green-digital co-specialisation that twin innovation requires. The critical mechanism is not the absence of path development activity, but the failure of cross-domain coupling — the incapacity of regional systems to stabilise green and digital knowledge trajectories within the same institutional and temporal space.

Table 3. Sufficiency Analysis

Conditions	Configurations					
	M1	M2	M3	M4	M5	M6
EKIA		⊗		●	●	
ICOLL		●	●	⊗	⊗	●
NEWB	⊗	⊗	⊗	●	⊗	●
FDI	⊗			⊗	●	●
UV	⊗		●	●	●	●
PS	●	●	●			⊗
RII	●	●	●	●	●	●
Consistency	0.825	0.809	0.829	0.858	0.875	0.849
PRI	0.381	0.417	0.478	0.482	0.561	0.529
Raw coverage	0.262	0.324	0.331	0.265	0.285	0.316
Unique coverage	0.002	0.003	0.007	0.002	0.029	0.027

Conditions	Configurations					
	M1	M2	M3	M4	M5	M6
Cases	DE32014 (0.5,0.35) DE32017 (0.54,0.4) DE42014 (0.50,0.2) DE82014 (0.50,0.1) DED2014 (0.51,0.4) DED2017 (0.57,0.4) DED2019 (0.63,1.0) DEE2014 (0.50,0.3) DEE2016 (0.5,0.18) DEG2011 (0.5,0.37) DEG2014 (0.5,0.37) DEG2016 (0.5,0.19) DEG2017 (0.5,0.23) DEG2019 (0.5,0.15) UKL2007 (0.5,0.32) UKN2014 (0.5,0.19)	AT12019 (0.6,0.17) AT22011 (0.66,0.4) AT22014 (0.5,0.36) AT22017 (0.53,1.0) AT22019 (0.62,0.4) AT32014 (0.5,0.14) AT32017 (0.60,0.1) AT32019 (0.71,0.1) BE32007 (0.5,0.15) BE32009 (0.6,0.25) BE32011 (0.59,0.2) BE32014 (0.5,0.27) BE32019 (0.50,0.2) DED2019 (0.52,1.0) DEE2014 (0.50,0.3) DEG2019 (0.5,0.15) FI1D2007 (0.8,0.4) FI1D2009 (0.8,0.2) FI1D2011 (0.7,0.4) FI1D2014 (0.65,1) FI1D2016 (0.5,0.4) FI1D2017 (0.6,0.4) FI1D2019 (0.61,1) FI182016 (0.50,1) UKC2014 (0.50,0.1) UKL2014 (0.50,0.3) UKL2017 (0.51,0.2) UKN2014 (0.5,0.19)	AT12019 (0.61,0.17) AT22011 (0.66,0.44) AT22014 (0.50,0.36) AT22017 (0.56,1.00) AT22019 (0.62,0.41) AT32014 (0.50,0.14) AT32016 (0.50,0.17) AT32017 (0.60,0.13) AT32019 (0.67,0.13) DE32011 (0.58,0.39) DE32019 (0.59,0.33) DEF2011 (0.51,0.13) FI1D2007 (0.81,0.4) FI1D2009 (0.80,0.2) FI1D2011 (0.7,0.46) FI1D2014 (0.65,1.0) FI1D2016 (0.51,0.4) FI1D2017 (0.6,0.39) FI1D2019 (0.61,1.0) FI182016 (0.50,1.00) FI192007 (0.54,0.32) FI192009 (0.52,0.50) FI192011 (0.57,1.00)	ES212007 (0.52,0.1) ES212014 (0.50,0.1) ES212016 (0.50,0.1) ITC12007 (0.73,0.38) ITC12009 (0.50,0.26) ITC12011 (0.65,0.35) ITC12014 (0.50,0.24) ITC12016 (0.50,0.22) ITH52007 (0.50,0.1) ITH52009 (0.50,0.1) ITH52011 (0.68,0.1) ITH52014 (0.50,0.1) ITI42007 (0.50,1.00) ITI42009 (0.50,0.44) ITI42011 (0.59,0.44)	AT12007 (0.54,0.33) SE112016 (0.6,0.49) SE112017 (0.7,0.32) SE232019 (0.6,1.00)	DE72014 (0.53,0.29) FR12011 (0.65,1.00) FR12016 (0.51,1.00) FR12017 (0.58,1.00) FR12019 (0.58,1.00) UKG2017 (0.51,0.28) UKH2007 (0.56,0.28) UKH2009 (0.51,0.32) UKH2014 (0.54,0.32) UKH2016 (0.50,0.33) UKH2017 (0.68,0.37) UKI2014 (0.50,0.28) UKI2016 (0.92,0.29) UKI2017 (0.93,0.18) UKI2019 (0.58,0.20) UKK2016 (0.76,1.00) UKM2017 (0.72,0.2)
RDP	-	Upgrading	Unrel. Div.	Creation Unrel. Div.	Importation Unrel. Div.	Creation Importation

Interpreted through the lens of regional path development theory (Grillitsch et al., 2018), the six configurations point to different empirical expressions of path-development potential, rather than to neatly separated path types. M1 stands apart as the only configuration without an identifiable path development mechanism, driven instead by policy support and innovation performance alone. The deviant consistency cases associated with this configuration display among the lowest outcome membership scores in the solution, suggesting that policy-induced activation without an underlying path-development mechanism produces at best transient alignment.

The remaining configurations reflect increasingly complex path dynamics, since their empirical profiles do not always map perfectly onto single theoretical mechanisms. This suggests that regional path development mechanisms may overlap in practice. M2 reflects an upgrading, where collaborative R&D activity and policy support sustain incremental transformation within existing knowledge structures, yet in the absence of entrepreneurial dynamism and structural complexity, the resulting twin innovation engagement remains weak and episodic. M3 points to an unrelated diversification, where cognitive variety activates cross-domain recombination in contexts of strong collaborative capacity — a more disruptive mechanism that nonetheless produces only partial conversion, as the absence of FDI and entrepreneurial activity limits the anchoring of new knowledge combinations within the regional system. M4 is best read as a hybrid configuration located at the intersection of creation and unrelated diversification, suggesting a route based on exogenous recombination and cognitively distant knowledge combinations. M5 is an importation-oriented configuration with elements of unrelated diversification, where external anchoring combines with structural complexity and cognitive variety.

M6 offers the clearest empirical signal of a realised twin innovation pathway. What distinguishes this configuration from the others is the density of its enabling conditions that co-occur in a mutually reinforcing constellation. This suggests that twin innovation does not emerge from the activation of any single path development mechanism in isolation, but requires the simultaneous stabilisation of endogenous agency, external knowledge anchoring, and cognitive variety within the same regional system. The concentration of high-membership cases among French regions points to the place-specific and institutionally thick character of this configuration.

Two findings deserve particular theoretical attention. First, path extension and branching — the mechanisms most widely documented in the EEG literature — are entirely absent from the sufficient configurations. This does not imply that these mechanisms are irrelevant to regional development in general, but that in this empirical setting they do not generate the cross-domain coupling required for twin innovation. Incremental, relatedness-based path dynamics are insufficient to bridge the cognitive and institutional distance between green and digital knowledge domains simultaneously. Second, the configurations that do support twin innovation are those involving the greatest degree of structural complexity, cognitive distance, and exogenous knowledge flows, pointing to the importance of hybrid path dynamics as drivers of twin innovation trajectories.

6 Conclusions

This paper has developed a configurational panel framework linking regional path development mechanisms to twin innovation outcome across European NUTS-2 regions over 2007–2019. By combining insights from Evolutionary Economic Geography and sustainability transitions research, and applying panel fuzzy-set Qualitative Comparative Analysis, the analysis has addressed a gap that neither tradition had systematically confronted: the conditions under which regional path development mechanisms support, constrain, or prevent the emergence of twin innovation trajectories.

Three findings stand out. First, twin innovation trajectories are rare, episodic, and geographically concentrated — not for lack of path-development activity, but because of the structural difficulty of coupling green and digital knowledge domains simultaneously within the same regional system over time. Second, incremental mechanisms such as path extension and branching — the most widely documented in the EEG literature — prove insufficient to generate this coupling. Only configurations combining endogenous recombination, external knowledge anchoring, and cognitive variety produce realised twin innovation trajectories, and even these remain temporally fragile. Third, the prevalence of deviant consistency cases reveals that path-development potential is widespread, but its conversion into twin innovation outcomes is systematically incomplete — a pattern that speaks to the dark side of regional path development and the limits of policy-induced activation without underlying structural depth.

These results carry direct implications for place-based innovation policy. Supporting twin innovation requires not only the development of regional capabilities in green and digital domains separately, but the deliberate cultivation of the cross-domain coupling conditions — relational, institutional, and cognitive — that allow these trajectories to stabilise and co-evolve within the same regional system over time.

7 Ethics declaration

This research did not require ethical approval, as it is based exclusively on publicly available secondary data and does not involve human participants, personal data, or sensitive information.

8 AI declaration

During the preparation of this work the authors used ChatGPT 4o in order to check grammar and spelling of the work and to improve the language style. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

References

Affolderbach, J., & Schulz, C. (2016). Mobile transitions: Exploring synergies for urban sustainability research. *Urban Studies*, *53*(9). https://doi.org/10.1177/0042098015583784

Asheim, B. T., Smith, H. L., & Oughton, C. (2011). Regional Innovation Systems: Theory, empirics and policy. *Regional Studies*, *45*(7), 875–891. https://doi.org/10.1080/00343404.2011.596701

Bachtrögler-Unger, Julia, Pierre-Alexandre, Boschma, Schwab, R., & Thomas. (2023). Technological capabilities and the twin transition in Europe: Opportunities for regional collaboration and economic cohesion. *MPRA*. https://doi.org/10.11586/2023017

Balland, P. A., Boschma, R., Crespo, J., & Rigby, D. L. (2019). Smart specialization policy in the European Union: relatedness, knowledge complexity and regional diversification. *Regional Studies*, *53*(9), 1252–1268. https://doi.org/10.1080/00343404.2018.1437900

Bhattacharya, P. (2023). How to build and analyze a panel data QCA model? A methodological demonstration of Garcia-Castro and Arino's panel data QCA model. *Methodological Innovations*, *16*(3), 265–275. https://doi.org/10.1177/20597991231179389

Blažek, J., Květoň, V., Baumgartinger-Seiringer, S., & Trippl, M. (2020). The dark side of regional industrial path development: towards a typology of trajectories of decline. *European Planning Studies*, *28*(8), 1455–1473. https://doi.org/10.1080/09654313.2019.1685466

Boschma, R. A., & Frenken, K. (2006). Why is economic geography not an evolutionary science? Towards an evolutionary economic geography. *Journal of Economic Geography*, *6*(3), 273–302. https://doi.org/10.1093/jeg/lbi022

Boschma, R., Coenen, L., Frenken, K., & Truffer, B. (2017). Towards a theory of regional diversification: combining insights from Evolutionary Economic Geography and Transition Studies. *Regional Studies*, *51*(1), 31–45. https://doi.org/10.1080/00343404.2016.1258460

Boschma, R., & Martin, R. (2010). The aims and scope of evolutionary economic geography. In *The Handbook of Evolutionary Economic Geography*. https://doi.org/10.4337/9781849806497.00007

Cantwell, J. (2010). Historical trends in international patterns of technological innovation. In *New Perspectives on the Late Victorian Economy*. https://doi.org/10.1017/cbo9780511660238.004

Coenen, L., Benneworth, P., & Truffer, B. (2012). Toward a spatial perspective on sustainability transitions. *Research Policy*, *41*(6). https://doi.org/10.1016/j.respol.2012.02.014

Cooke, P. (2013). Complex Adaptive Innovation Systems: Relatedness and transversality in the evolving region. In *Complex Adaptive Innovation Systems: Relatedness and Transversality in the Evolving Region*. https://doi.org/10.4324/9780203126615

Fazio, G., Maioli, S., & Rujimora, N. (2024). The twin innovation transitions of European regions. *Regional Studies*. https://doi.org/10.1080/00343404.2024.2309176

Fuenfschilling, L., & Truffer, B. (2014). The structuration of socio-technical regimes - Conceptual foundations from institutional theory. *Research Policy*, *43*(4). https://doi.org/10.1016/j.respol.2013.10.010

Garcia-Castro, R., & Ariño, M. A. (2016). A General Approach to Panel Data Set-Theoretic Research. *Journal of Advances in Management Sciences & Information Systems*, *2*, 63–76. https://doi.org/10.6000/2371-1647.2016.02.06

Gibbs, D., & O'Neill, K. (2017). Future green economies and regional development: a research agenda. *Regional Studies*, *51*(1). https://doi.org/10.1080/00343404.2016.1255719

Grillitsch, M., Asheim, B., & Trippl, M. (2018). Unrelated knowledge combinations: The unexplored potential for regional industrial path development. *Cambridge Journal of Regions, Economy and Society*, *11*(2), 257–274. https://doi.org/10.1093/cjres/rsy012

Hassink, R., Isaksen, A., & Trippl, M. (2019). Towards a comprehensive understanding of new regional industrial path development. *Regional Studies*, *53*(11), 1636–1645. https://doi.org/10.1080/00343404.2019.1566704

Hodson, M., Geels, F. W., & McMeekin, A. (2017). Reconfiguring urban sustainability transitions, analysing multiplicity. *Sustainability (Switzerland)*, *9*(2). https://doi.org/10.3390/su9020299

Isaksen. (2020). Path Development in Different Regional Innovation Systems: A Conceptual Analysis. In *Innovation Drivers and Regional Innovation Strategies*. https://doi.org/10.4324/9781315671475-12

Karlsson, C. (2008). Handbook of research on cluster theory. In *Handbook of Research on Cluster Theory*. https://doi.org/10.4337/9781848442849

Köhler, J., Geels, F. W., Kern, F., Markard, J., Onsongo, E., Wieczorek, A., Alkemade, F., Avelino, F., Bergek, A., Boons, F., Fünfschilling, L., Hess, D., Holtz, G., Hyysalo, S., Jenkins, K., Kivimaa, P., Martiskainen, M., McMeekin, A., Mühlemeier, M. S., … Wells, P. (2019). An agenda for sustainability transitions research: State of the art and future directions. *Environmental Innovation and Societal Transitions*, *31*. https://doi.org/10.1016/j.eist.2019.01.004

Loorbach, D., Frantzeskaki, N., & Avelino, F. (2017). Sustainability Transitions Research: Transforming Science and Practice for Societal Change. In *Annual Review of Environment and Resources* (Vol. 42). https://doi.org/10.1146/annurev-environ-102014-021340

Mäkitie, T., Hanson, J., Damman, S., & Wardeberg, M. (2023). Digital innovation's contribution to sustainability transitions. *Technology in Society*, *73*. https://doi.org/10.1016/j.techsoc.2023.102255

Martin, R., & Sunley, P. (2006). Path dependence and regional economic evolution. *Journal of Economic Geography*, *6*(4), 395–437. https://doi.org/10.1093/jeg/lbl012

Montresor, S., & Vezzani, A. (2023). Digital technologies and eco-innovation. Evidence of the twin transition from Italian firms. *Industry and Innovation*, *30*(7), 766–800. https://doi.org/10.1080/13662716.2023.2213179

Oana, I.-E., & Schneider, C. Q. (2018). SetMethods: an Add-on R Package for Advanced QCA. *The R Journal*. http://compasss.org/software.htm

Pinheiro, F. L., Balland, P. A., Boschma, R., & Hartmann, D. (2022). The dark side of the geography of innovation: relatedness, complexity and regional inequality in Europe. *Regional Studies*. https://doi.org/10.1080/00343404.2022.2106362

Ragin, C. C. (2000). *Fuzzy-Set Social Science*. The University of Chicago Press.

Ragin, C. C. (2008). Redesigning social inquiry: Fuzzy sets and beyond. In *University of Chicago Press*. The University of Chicago Press. https://doi.org/10.7208/chicago/9780226702797.001.0001

Rihoux, B., & Ragin, C. C. (2009). Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques. In Sage (Ed.), *Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques*. https://doi.org/10.4135/9781452226569

Strambach, S., & Pflitsch, G. (2020). Transition topology: Capturing institutional dynamics in regional development paths to sustainability. *Research Policy*, *49*(7). https://doi.org/10.1016/j.respol.2020.104006

Truffer, B., & Coenen, L. (2012). Environmental Innovation and Sustainability Transitions in Regional Studies. *Regional Studies*, *46*(1). https://doi.org/10.1080/00343404.2012.646164

Vertova, G. (2001). National technological specialisation and the highest technological opportunities historically. *Technovation*, *21*, 605–612. https://doi.org/10.1016/S0166-4972(01)00034-7

Woodside, A. G. (2013). Moving beyond multiple regression analysis to algorithms: Calling for adoption of a paradigm shift from symmetric to asymmetric thinking in data analysis and crafting theory. In *Journal of Business Research* (Vol. 66, Number 4). https://doi.org/10.1016/j.jbusres.2012.12.021

Knowledge Management and Artificial Intelligence: Governance and Liability in Business Management

Claudio D'Alonzo
Catholic University Our Lady of Good Counsel, Tirana, Albania

Abstract

The purpose of this paper is to analyse the role of artificial intelligence in corporate management. Since the topic does not concern only the digital sector, the research is conducted from a legal perspective. Consequently, the methodological approach adopted considers both the existing legal framework and the relevant legal doctrine; however, it does not seem possible to make reference to case law. As a result, this article has a theoretical character and adopts a conceptual approach to the subject matter.

In Italy, not only does the AI Act apply, but also Law No. 132 of 23 September 2025, which emphasises the strategic importance of artificial intelligence for the national economy. The foregoing considerations suggest that companies are required to undertake a profound review of their organisational and operational models.

The findings highlight that it is possible to identify three distinct levels of interaction between corporate law and artificial intelligence tools. The first concerns support for the decision-making process of the management body. The second relates to the use of artificial intelligence to improve dialogue with shareholders and third parties in general. Finally, artificial intelligence is considered a tool for the external assessment of a company's functioning.

One of the specific issues that arises concerns the need to assess the extent to which artificial intelligence can be integrated into the organisation of a company. This raises the question of whether the use of AI is necessary to such an extent that failure to adopt it would result in inadequate corporate structures.

Finally, the study highlights that the issue of adequate organisational structures must enable directors to continuously monitor the adopted artificial intelligence system from the moment of its selection through to its use, also in order to prevent undesirable outcomes. Clearly, AI does not replace directors, but it serves as a support tool for them.

Keywords: artificial intelligence, knowledge management, corporate management, organisational structures, existing legislation

Paper type: Academic Research Paper

1 Introduction

Knowledge is one of an organisation's most valuable resources and forms the basis for the creation and development of the organisation's innovative capacity. In company law, knowledge plays a key role as a fundamental element of organisational structures and decision-making processes, directly impacting the obligations of proper management and the liability of directors. In this context, a pivotal role is played by new technologies (Nascimento et al., 2021), particularly artificial intelligence (AI). AI appears to be a fundamental element of knowledge management (KM) thanks to improved information collection, development, and dissemination, as well as its efficient implementation within organisations (He and Burger-Helmchen, 2025).

The recent increase in popularity of AI and its diffusion across various sectors has radically changed the role of KM, which has been transformed into a complex discipline aimed at improving decision-making and increasing value creation (Sanzogni et al., 2017; Al Mansoori et al., 2020; Leoni et al., 2024). To achieve this, it is essential to establish a solid relationship between existing knowledge (Iqbal, 2023) and the AI tools that support KM practices (Arias-Pérez et al., 2022).

The purpose of this paper is to analyse the role of AI in corporate management. Since the topic does not concern only the digital sector, the research is conducted from a legal perspective. Consequently, the methodological approach adopted takes into account both the existing legal framework and the relevant legal doctrine; however, it does not seem possible to make reference to case law. As a result, this article has a theoretical character and adopts a conceptual approach to the subject matter. The findings show that AI is part of the organisation of a company and that directors are required to monitor the adopted AI system constantly, also in order to prevent undesirable outcomes.

The paper is organised as follows. Section 2 deals with KM in a business. Section 3 presents the existing regulatory framework. Sections 4, 5, and 6 address the skills of directors and the integration of AI in a business. Section 7 concerns the liability of directors. The last part presents the conclusions.

2 KM in a Business

The convergence of AI and KM is not a new concept in the scientific field nor in the business world; in fact, modern businesses are investing significantly in KM. In business law, KM can be understood as the set of tools, techniques, and strategies for storing, analysing, organising, improving, and sharing entrepreneurial experience (Groff and Jones, 2003). The evolution of many businesses in recent times is closely linked to the increasingly widespread use of AI (Abriani, 2022), which has rapidly spread across various sectors such as healthcare and finance. Beyond the risks that may arise from the use of this technology (for example, biases and discrimination), there is a need to assess the consequences for companies that use AI to carry out their business.

Intelligent knowledge, i.e., the synergistic integration of human cognitive capabilities and AI-based tools, is transforming business law, both nationally and internationally. However, KM remains a constant challenge due to multiple factors, such as the risk of losing tacit knowledge or information overload (Jarrahi et al., 2023). For this reason, companies are required to transform continuously; as a consequence, innovation becomes a goal, and not just a way of thinking, to be absorbed in terms of culture, best practices, and initiatives.

The relationship between innovation and law has never been simple, but AI has made this debate even more urgent and complex. Business law is at a crossroads: on the one hand, it must adapt to the new operating methods introduced by intelligent technologies; on the other, it must remain linked to the principles that guarantee the integrity of the legal system and the protection of collective interests.

In an era of digital acceleration and AI-driven transformation, businesses must not focus exclusively on technology as the primary driver of innovation; this is because innovation is not just a technological issue, but above all a cultural and human one. To achieve a true advantage, businesses are required to evolve towards intelligent systems, where organisation is viewed as a living organism capable of learning and adapting. From this perspective, knowledge is not a static asset (stored in a database), but a dynamic flow that constantly feeds decision-making processes.

3 The Existing Regulatory Framework

The existing regulatory framework for AI has significant impacts on businesses and their organisations. Therefore, it seems useful to examine the role AI plays in the life of a business, considering that current legislation, both at the European and national levels, promotes the correct, responsible, and transparent use of this tool; all from a perspective that places the human being at the centre.

As is well known, the increasingly prominent role played by AI has led the European lawmaker to enact the European Artificial Intelligence Act (AI Act). The scope for national discretion left by European legislation has been promptly incorporated into Italian law through Law No. 132 of 23 September 2025, in which information, and consequently the management and dissemination of knowledge, play a fundamental role. Italy is the first EU Member State to have adopted a national legislative instrument dedicated exclusively to the regulation of AI, anticipating the operational implementation of the AI Act (Sottoriva and Vetrugno, 2025). In this way, the Italian law stands out for promoting a national strategy capable of reconciling technical, ethical and social requirements (Pesucci, 2026). The legislation issued sets out key principles such as transparency, data protection, and security (Raganelli, 2025). In this context, with a view to initiating a transformation of corporate compliance systems and internal governance (Sottoriva and Vetrugno, 2025), Italian law provides, amongst other things, that the State and other public authorities shall promote the development and use of AI as a tool to improve human-machine interaction.

The law provides detailed and sector-specific provisions that build on the existing regulatory framework, particularly regarding compliance, transparency, accountability, and governance. Consequently, businesses are required to modify their organisational models both to integrate AI systems into decision-making, management, and production processes, and to address the risks arising from the use of such systems, avoiding liability for non-compliance with current legislation.

4 The Skills of Directors

The above considerations lead to reflecting, albeit briefly, on the impact that AI has on corporate governance. Indeed, AI directly influences the way management functions are exercised. First, it seems possible to state that AI and KM, together, may enable a company to improve its results, increase its profitability, and enhance its market competitiveness. It follows that AI serves, first and foremost, as a tool to support the organisation of a company and make it more competitive on the market.

The first perspective from which the interaction between corporate management and AI may be examined concerns the possibility of providing support in the identification of potential directors (Rimini, 2024). In this

regard, it is worth noting that there has been a past attempt to automate the administrative body through the so-called DAO Project (Picciau, 2021), based on the ambitious and provocative idea that company management could be more efficient if entrusted to intelligent computers rather than to humans (Rimini, 2024; Abriani, 2022). Although it was not successful and it seems unlikely that AI will lead to the disappearance of a company's management body or make the delegation of functions to one or more directors superfluous (Benedetti, 2025), the proposal highlights the attempt to innovate the principles of corporate governance through a functional replacement of decision-making roles.

Under current legislation, even though directors are not required to have high levels of technological expertise, it seems clear that the use of AI requires at least a basic understanding of such systems, as well as the ability to understand their limitations, risks, and potential. Therefore, the specific technical skills of individual candidates should be adequately assessed when appointing directors. Consequently, it seems appropriate for a company's administrative body to consist of at least one member with IT expertise, who can contribute to any preliminary discussions prior to a specific decision and broaden the discussion within the board (Ruffolo and Amidei, 2025; Paccoia, 2025; Enriques and Zorzi, 2023; Pacileo, 2022). This solution does not appear illogical, being it normal for the board of directors of a medium-large company to be composed of several different professional figures with specific skills. From this perspective, it is essential that directors undergo continuous training through appropriate programs that focus on digital innovation and the risks associated with it (Mäntymäki, 2023).

In light of the above, it is necessary to consider how human intelligence and AI might interact. The adoption of AI-based tools by businesses calls for a re-examination of governance issues, whilst ensuring compliance not only with current legislation but also with ethical standards (Rahul, 2025).

5 The Integration of AI into a Business

It is possible to identify three distinct levels of interaction between corporate law and AI tools; this is because AI may be used primarily to support the activity carried out (De Nadai, 2024; Montagnani, 2020). AI has long been present in the life of a company, inserting itself into the decision-making processes and influencing the choices of the management bodies (Enriques and Zorzi, 2023). Although the current legislation is aimed at humans (Lucarelli Tonini, 2023), given that AI systems are not legal persons, it is necessary to recognise that management decisions no longer depend exclusively on human will (Rimini, 2024). In fact, the use of AI may support both the day-to-day management of a company and the guidance and evaluation phase of the administrative body (Abriani, 2022; Benedetti, 2025). Examples include the examination of complex contracts, financial documents, and other critical information quickly and with a high degree of reliability. This does not mean predicting the outcomes of every operation (Enriques and Zorzi, 2023), but understanding the risk factors that could expose a business to potential uncertainties, such as legal, financial, or reputational liability. It follows that the integration of AI and business requires a profound change in the way directors fulfil their duties.

The legislative introduction of transparency and traceability criteria for AI systems used by companies leads to new disclosure requirements. This aspect is particularly relevant in all companies operating in the AI, cybersecurity, and instrumental technologies sectors. In this way, digitalisation paves the way for Italy's technological strengthening, at least within Europe. The duty of transparency falls on managing directors, who must inform the board of directors as well as shareholders. It seems possible to argue that the general obligation to act in an informed manner requires an assessment not only of the choices made through the use of AI tools, but also of the specific characteristics of the system used (Rimini, 2024); of course, the duty to provide information varies with the model used and must be inevitably fulfilled more rigorously where the chosen system is provided or managed by a third party.

However, the question arises as to whether a company's governing bodies must necessarily adhere to the instructions received or whether they can deviate from them. The most preferable solution appears to be to adopt a comply-or-explain approach, meaning that when the AI's instructions are not followed it is necessary to specify why that happened (Picciau, 2021). Therefore, every decision made by the governing body must be not only transparent and accountable (Štrukelj and Gajšt, 2025), but must also allow for an assessment of the logical path followed.

The second level of connection between business and AI concerns the use of AI to improve the organisation and functioning of a company's governing bodies, to ensure compliance with legal and statutory obligations, and to foster dialogue with shareholders and third parties in general (Abriani, 2022). On this point, it seems worth noting that KM and its transmission within a company are sensitive issues, as it is not always possible to share information among the various bodies that make up a company. Although a continuous flow of information among those responsible for managing and supervising the company is normal and it is natural that their rotation

is followed by the transfer of information, it is not necessarily the case that this information should be made available to the shareholders. However, the possibility that AI could be used as a tool to strengthen the dialogue between directors and shareholders, reducing information asymmetries between these two bodies, seems not remote. Likewise, AI could be helpful in assessing the impact that any changes to the administrative and financial rights of company shares may have on ownership structures (Montagnani, 2021). This creates a new scenario in which companies are required not only to evaluate the cost-benefit ratio in day-to-day management and in their relationships with shareholders, but also to address new challenges, including liability arising from the use of these new digital tools, known as Corporate Digital Responsibility (Wade, 2020; Lobschat et al., 2021; Abriani and Schneider, 2020; Valiante et al., 2023).

Finally, AI is being considered as a tool for assessing a company's performance from an external perspective. This approach is justified by the fact that improved information flows between directors and shareholders may inevitably be followed by increased communications directed at third parties (Abriani, 2022).

6 AI and Business' Structure

The integration of AI within a business raises fundamental questions regarding the duties assigned to the administrative body. In Italy, pursuant to Article 2086 of the Italian Civil Code, directors are required to *"establish an organisational, administrative, and accounting structure appropriate to the nature and size of the company, also with a view to the timely identification of business crises and the loss of business continuity"*.

One of the main questions posed by the current legislative and social landscape concerns the role that AI solutions can play within an appropriate organisational, administrative, and accounting structure. It seems undeniable that the assistance offered by AI tools favours the inclusion of such systems within the organisation of a company (Ruffolo and Amidei, 2025).

Therefore, not only does AI become an integral part of the organisational, administrative, and accounting structure, but it also contributes to compliance with the law and internal regulations (De Nadai, 2024; Enriques and Zorzi, 2023). In particular, with regard to this last aspect, the inclusion of AI in a company's organisation may help identify the concrete possibilities for business continuity and the effective sustainability of its debts.

Obviously, the integration of AI into a business organisation raises the question of whether its use is so necessary that failure to adopt it would result in the inadequacy of the corporate structure (Abriani, 2022; Enriques and Zorzi, 2023). If this were the case, this would have implications for directors' liability. It seems safe to say that the answer to this question is not unequivocal, but varies based on the nature of the business performed and the size of the company (Benedetti, 2025; Montagnani, 2020). The function of AI is to ensure compliance with the law and to support directors in fulfilling their duties, but there is no obligation to use it. However, where this occurs, it is essential that the administrative body use a reliable and transparent AI system (Bruner, 2021), without prejudice to the need to evaluate the response provided and decide whether or not to carry out a specific operation. The arrangements in place must allow directors to constantly monitor the AI system adopted from the moment of its selection until its use, also in order to prevent undesirable outcomes (Pacileo, 2022).

In any case, it seems essential to diversify the administrative bodies by establishing specific committees (De Nadai, 2024) which, based on the nature and size of a company, adopt procedures suitable for the appropriate use of AI and the management of the risks associated with it, such as regulatory, technical and ethical ones (Abriani, 2022; Moerel, 2020); in this perspective, the creation of a technology committee (Rimini, 2024), possibly composed of independent directors (Enriques and Zorzi, 2023), which deals with the choice of AI systems and their monitoring, could be useful (Pacileo, 2022); although a committee of this type plays a strategic role within a company, it remains understood that it is not vested with management powers and that every decision is up to the board of directors as a whole.

In addition to the creation of a specific committee, it would also be useful to appoint a data protection officer who could play a central role in AI governance to ensure compliance in the handling of personal data.

It follows that administrative bodies should not only be made up of adequately experienced and competent figures in AI, but should also recruit cybersecurity professionals and ensure that there is a flow of information with them. AI must be embedded within the organisational structure of a company, so that it can act as a powerful cognitive technology and adequately manage all the information at its disposal.

7 Directors' Liability

It has been said that AI has become an integral part of businesses and their decision-making processes. Decisions made by the board of directors are characterised by a high degree of discretion, as they are the final stage of a complex process that requires the evaluation of a series of factors that are not easily measurable.

Although the use of AI may be necessary in some cases, the board of directors must not be passive, relying entirely on its decisions without any critical evaluation. Instead, it is always necessary to maintain an active role and engage directly in the decision-making processes. Otherwise, directors could be held liable for failing to act diligently and for abdicating their role in favour of the results obtained by the adopted AI system. The foregoing considerations lead to the belief that the role of AI in the organisation of a company is similar to that of external consultants who provide opinions or recommendations (Ruffolo and Amidei, 2025); AI is an aid for directors, who must evaluate whether or not to proceed with a specific transaction, taking into due consideration that they may incur liability towards the company if they do not use the required diligence.

It remains clear that decisions taken do not need to be justified and that the principle of the business judgment rule applies, whose purpose is to prevent the transfer of business risk from shareholders to directors (Pacileo, 2022). On this point, it may be useful to note that directors must be able to prove that they acted in an informed manner, in good faith, and in the company's best interest. However, where a decision is made using AI tools, only in the presence of specific information on the AI system used, its level of transparency, and the role it played in making certain decisions will it be possible to exclude directors' liability, even where the transaction had a negative outcome for the company (Ruffolo and Amidei, 2025; Montagnani, 2020; Rimini, 2024).

The issue of liability appears to intersect with that of administrative, organisational, and accounting structures (Pacileo, 2022; Perrino, 2025). It seems fair to say that the duty to create an organisational structure adequate for the activity carried out cannot be subject to stringent judicial oversight and, consequently, is covered by the business judgment rule. It follows that the violation of the obligation to establish adequate organisational, administrative, and accounting structures cannot be an independent cause for liability of the administrative body, but only a contributing cause to be examined together with other charges (Paccoia, 2025). With reference to the issue of AI, in order to benefit from the protection of the business judgment rule, directors will have to ensure that AI-based structures are adequate also in terms of general reliability (Enriques and Zorzi, 2023). From this perspective, it seems necessary to document the characteristics of the technology used and the instructions given in written and comprehensible form. In other words, the administrative body is responsible for disclosing the AI system used and its operating methods. The need to track AI decisions not only appears appropriate from a corporate law perspective, but also complies with the provisions of the applicable legislation, which establishes transparency, traceability, and clarity (Ruffolo, 2025).

8 Conclusions

As mentioned, AI is a tool that enables the transmission of various branches of knowledge and may become essential both for facilitating decision-making processes and for providing decision makers with a broader range of possible solutions. The use of AI in corporate law strikes a delicate balance between opportunities and pitfalls. AI is a formidable innovation tool, capable of optimising resources, simplifying decision-making processes, and increasing market competitiveness. The automation of administrative functions, predictive analytics, and the ability to access automated systems are strategic advantages for businesses. AI also influences a company's governance structure, shifting the focus of decision-making toward automated tools that require constant human oversight.

The competitive skill of a modern company lies, first and foremost, in its ability to manage and govern knowledge. KM directly impacts the innovative capacity of a company as an organisation, considering that innovation is not solely linked to the adoption of advanced digital tools, but also (and above all) to the cognitive dynamics that form its structural foundation.

AI can be an integral part of a company's organisation, albeit within certain limits. First, AI can assist the board of directors in day-to-day management, but it cannot replace human judgment in strategic decisions. In other words, AI does not replace directors, but rather supports them. This means that directors must not take the results obtained for granted, trusting in a high probability of accuracy and reliability (Pacileo, 2022). It follows that despite the entry of new technologies into corporate life, its management still requires a human-centric approach in which AI is a tool and not a solution (Sacco Ginevri, 2021). This interpretation is justified by the fact that computers and information systems, however efficient, lack typically human characteristics such as flexibility, adaptability, and communication skills (Picciau, 2021). What has been said so far is in line with current national and European legislation, which emphasises that AI's role is to assist humans in decision-making processes and not to replace them. The crucial role entrusted to directors and their specific skills also in terms of technological innovation are the key to avoiding the total transfer of business management to AI, including the establishment of organisational structures.

9 Ethics Declaration

The Author declares that ethical clearance was not required.

10 AI Declaration

The Author declares that AI tools were not used for the creation of the paper

References

Nascimento, L.d.S., Reichert, F.M., Janissek-Muniz, R. and Zawislak, P.A. (2021) Dynamic interactions among knowledge management, strategic foresight and emerging technologies, *Journal of Knowledge Management*, Vol. 25, No. 2, pp. 275-297.

He X., Burger-Helmchen T. (2025) Evolving Knowledge Management: Artificial Intelligence and the Dynamics of Social Interactions, *Ieee Engineering Management Review*, Vol. 53, No. 5, 2025, pp. 215-231.

Sanzogni, L., Guzman, G. and Busch, P. (2017) Artificial intelligence and knowledge management: questioning the tacit dimension, *Prometheus*, Vol. 35, No. 1, pp. 37–56.

Al Mansoori, S., Salloum, S.A. and Shaalan, K., (2020) "The impact of artificial intelligence and information technologies on the efficiency of knowledge management at modern organizations: a systematic review", *Recent advances in intelligent systems and smart applications*, pp.163-182.

Leoni, L., Gueli, G., Ardolino, M., Panizzon, M. and Gupta, S. (2024) AI-empowered KM processes for decision-making: empirical evidence from worldwide organisations, *Journal of Knowledge Management*, Vol. 28, No. 11, pp. 320-347.

Iqbal, U. (2023) From Knowledge Management to Intelligence Engineering - A practical approach to building AI inside the law-firm using open-source Large Language Models, *Proceedings of the Third International Workshop on Artificial Intelligence and Intelligent Assistance for Legal Professionals in the Digital Workplace*, pp. 49-57.

Arias-Pérez, J. and Vélez-Jaramillo, J. (2022) "Understanding knowledge hiding under technological turbulence caused by artificial intelligence and robotics", *Journal of Knowledge Management*, Vol.26 No.6, pp.1476 1491.

Groff, T. and Jones, T. (2003) *Introduction to Knowledge Management: KM in Business*, London.

Abriani, N. (2022) L'impatto dell'intelligenza artificiale sulla governance societaria: sostenibilità e creazione di valore nel lungo termine, *Contratto e Impresa Europa*, pp. 89-125.

Jarrahi, M. H., Askay, D., Eshraghi, A., & Smith, P. (2023). Artificial intelligence and knowledge management: A partnership between human and AI. Business Horizons, 66(1), pp. 87–99.

Mäntymäki, M., Minkkinen, M., Birkstedt T., Viljanen M. (2023) Putting AI Ethics into Practice: The Hourglass Model of Organizational AI Governance, arXiv:2206.00335v2

Sottoriva, C. and Vetrugno, G. (2025) La legge sull'Intelligenza Artificiale: inquadramento normativo e ricadute sul sistema giuridico-economico nazionale, *Ius Societario*, pp. 1-7.

Pesucci, G. (2026) AI e sostenibilità: dalla tecnica alla governance, *Ilcaso.it*, pp. 1-16, https://www.ilcaso.it/uploads/allegati_saggi/simone_pesucci-18-01-2026-0e5e6.pdf

Raganelli, B. (2025) Crisi, sostenibilità sociale e debt advice. *Rivista di diritto bancario*, p. 371-398.

Rimini, E. (2024) Intelligenza artificiale e doveri degli amministratori, *Rivista di corporate governance*, pp. 87-105.

Picciau, C. (2021) The (Un)Predictable Impact of Technology on Corporate Governance, *Hastings Business Law Journal*, vol. 17, n. 1, pp. 67-136.

Benedetti, L. (2025) Il paradosso dell'AI come strumento (necessario) di adempimento e come causa di violazione dei doveri degli amministratori nell'artificial intelligence based corporate governance, *Giurisprudenza commerciale*, pp. 632-656.

Ruffolo, U. and Amidei, A. (2025) *Diritto dell'intelligenza artificiale*, Roma.

Paccoia, M. (2025) *Gli assetti adeguati* ex *art. 2086 c.c. Il punto di vista del giurista*, Torino.

Enriques, L. and Zorzi, A. (2023) "Intelligenza artificiale e responsabilità degli amministratori", *RDS*, pp. 13-46.

Pacileo, F. (2022) "Scelte d'impresa" e doveri degli amministratori nell'impiego dell'intelligenza artificiale", *RDS*, pp. 539-578.

Rahul, G. (2025) AI in corporate governance: legal, ethical, and fiduciary challenges in the age of artificial intelligence, *International Journal of Legal Studies and Social Sciences*, vol. 3, pp. 314-322.

De Nadai, M. (2024) L'impiego degli strumenti di intelligenza artificiale da parte degli amministratori: rischi e responsabilità, *Ristrutturazioni aziendali*, 2 October, pp. 1-13.

Montagnani, M. L. (2020) Intelligenza artificiale e governance della "nuova" grande impresa azionaria: potenzialità e questioni endoconsiliari, *Rivista delle società*, pp. 1003-1026.

Lucarelli Tonini, L.M. (2023) L'IA tra trasparenza e nuovi profili di responsabilità: la nuova proposta di "IA Liability Directive", *Il diritto dell'informazione e dell'informatica*, 2023, pp. 327-362.

Kahneman, D, Rosenfield, A.M., Gandhi, L., Blaser, T. (2016) Noise: How to Overcome the High, Hidden Cost of Inconsistent Decision Making, *Harvard Business Review*, https://hbr.org/2016/10/noise.

Štrukelj, T., Gajšt, N. (2025) From Routine to Intuition: AI's Role in Organization's Governance and Management Decision-Making, *9th FEB International Scientific Conference: Sustainable Management in the Age of ESG and AI: Navigating Challenges and Opportunities*, pp. 892-899.

Montagnani, M.L. (2021) *Il ruolo dell'intelligenza artificiale nel funzionamento del consiglio di amministrazione delle società per azioni*, Milano.

Wade, M. (2020) Corporate Responsibility in the Digital Era, MIT Sloan Management Review, https://sloanreview.mit.edu/article/corporate-responsibility-in-the-digital-era/

Lobschat, L., Mueller, B., Eggers, F., Brandimarte, L., Diefenbach, S., Kroschke, M., Wirtz, J. (2021) Corporate Digital Responsibility, *Journal of Business Research*, 2021, 122, pp. 875-888.

Abriani, N., Schneider, G. (2020) Il diritto societario incontra il diritto dell'informazione. IT, Corporate governance e Corporate Social Responsibility, *Rivista delle società*, pp. 1326-1397.

Valiante, D., Lamandini, M., Ramos Muñoz, D. (2023) The European Strategy on Digital Finance and Its Interplay with Capital Markets Integration in the EU, Cambridge.

Bruner, C.M. (2021) Artificially Intelligent Boards and the Future of Delaware Corporate Law, University of Georgia School of Law Legal Studies Research Paper No. 2021-23, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3928237

Moerel, L. (2020) Reflections on the Impact of the Digital Revolution on Corporate Governance of Listed Companies, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3519872

Locascio Aliberti, E. (2025) Appunti sul Comitato per la Sostenibilità (e sull'opportuno accorpamento al comitato controllo e rischi, *Giurisprudenza commerciale*, pp. 559-588.

Perrino, M. (2025) Adeguatezza organizzativa, amministrativa e contabile e doveri dell'imprenditore. Declinazioni dell'adeguatezza come "dovere", *RDS*, pp. 29-52.

Ruffolo, U. (2025) L'AI Act e l'evoluzione dell'AI, *AI Act. La regolamentazione europea dell'intelligenza artificiale*, Roma.

Sacco Ginevri, A. (2021) Ancora su intelligenza artificiale e corporate governance, *Rivista trimestrale di diritto dell'economia*, pp. 343-352.

Montagnani, M. L. (2020) Flussi informativi e doveri degli amministratori di società per azioni ai tempi dell'intelligenza artificiale, *Persona e mercato*, pp. 86-105.

Picciau, C. (2022) Intelligenza artificiale, scelte gestorie e organizzazione delle società per azioni, *Intelligenza artificiale e diritto: una rivoluzione?*, Bologna.

Between Heritage and Change: Value-Oriented Digital Transformation in European Theatres

Dóra Horváth
Corvinus University of Budapest, Budapest, Hungary

Abstract

This study examines how theatres interpret, govern, and strategically integrate digital technologies and knowledge practices in the context of mission-driven value creation. While digital transformation is often associated with efficiency gains and technological change, less attention has been paid to how digital knowledge is embedded in organizational practices and contributes to cultural and social value in non-profit contexts. Drawing on 28 semi-structured interviews with theatre professionals across Europe, the study applies an inductive qualitative approach inspired by grounded theory to explore how digitalization unfolds in practice. The findings show that digitalization is primarily embedded in operational and support activities, where digital tools function as knowledge infrastructures that enhance communication, coordination, and audience engagement. Rather than being driven mainly by efficiency, these practices contribute to public value by extending access, sustaining relationships, and reinforcing theatres' cultural missions. At the same time, digital transformation remains incremental and selective. More advanced forms of transformation - defined as the strategic integration of digital knowledge into governance and value creation - are limited, as digital practices are typically embedded within existing institutional frameworks shaped by legacy structures and public value orientations. Digital experimentation supports organizational learning but remains largely project-based and weakly institutionalized. Overall, the study shows that digital transformation in theatres primarily reinforces existing value creation logics rather than redefining them. It contributes to the literature by conceptualizing digital transformation as a knowledge-based, value-oriented process shaped by governance, institutional logics, and organizational capacities, offering insights into how mission-driven organizations navigate digital change beyond market-based contexts.

Keywords: cultural organizations, non-profit organizations, cultural heritage, mission-driven, digital transformation

Paper type: Academic Research Paper

1 Introduction

In the current phase of digital transformation, the central challenge for organizations is not simply whether digital technologies improve efficiency, but how efficiency gains are integrated into broader processes of value creation and organizational practice shaped by digital knowledge. In this sense, digital technologies increasingly influence how knowledge is integrated into strategic decision-making, accountability, and relationships with stakeholders, thereby altering organizational priorities and forms of legitimacy (Matt *et al.*, 2015; Vial, 2019; Wessel *et al.*, 2021). These dynamics are particularly consequential for non-profit, public organizations, where digital transformation intersects directly with mission fulfilment, public responsibility, and the production of cultural and social value rather than market-based returns (Bakhshi and Throsby, 2012). In such contexts, digital systems do not operate as neutral infrastructures but as structuring forces that shape governance arrangements and organizational identities (Hanisch *et al.*, 2023), while their integration remains conditioned by historically embedded institutional logics and path-dependent practices (Johnson, 2007).

Theatre organizations represent a distinctive case within this broader transformation. As historically embedded cultural institutions, theatres are deeply rooted in traditions of live presence, artistic autonomy, and cultural heritage. They operate predominantly as non-profit or publicly supported organizations, shaped by strong institutional legacies and path-dependent routines that prioritize continuity, legitimacy, and public responsibility (Bonet and Schargorodsky, 2018; Johnson, 2007; Schreyögg *et al.*, 2011). At the same time, theatres are increasingly exposed to digital pressures that affect the entire cultural value chain, from artistic creation and production to dissemination, audience engagement, and organizational coordination (Ford and Mandviwalla, 2020; De Voldere *et al.*, 2017). This creates a structural tension between digital transformation as a force of organizational reconfiguration and theatres as tradition-oriented, mission-driven organizations. Digital technologies such as audience data systems, digital archives, streaming platforms, and hybrid performance formats actively shape organizational learning, redefine relationships with audiences, and introduce new forms of knowledge-based coordination (Saarikko *et al.*, 2020). However, their integration is mediated by institutional logics centered on authenticity, embodied co-presence, and cultural continuity, which constrain radical or

disruptive change (Lampel *et al.*, 2000). From a theoretical perspective, digital transformation in theatres unfolds primarily through selective and negotiated integration. Digital knowledge systems reshape value creation beyond profit by extending access, supporting cultural preservation, and enabling more inclusive and knowledge-informed governance, while remaining embedded in historic missions and public value logics (Ford and Mandviwalla, 2020; Hawthorne, 2023). Despite these contributions, existing research offers limited empirical insight into how digital transformation is strategically governed and knowledge is mobilized for value-oriented outcomes in historically embedded, mission-driven cultural organizations such as theatres, particularly beyond isolated digital practices or technology-specific interventions.

To address this gap, the study adopts a qualitative research design based on semi-structured interviews with theatre leaders and key professionals across Europe. A total of 28 in-depth interviews were conducted with executive directors, artistic directors, marketing and communication managers, financial leaders, innovation managers, artists, researchers, and digital service providers. The interviews examined how theatres interpret, govern, and strategically integrate digital technologies and knowledge practices to reconcile efficiency gains with mission-driven value creation, public responsibility, and long-term cultural relevance. The interview data were analysed using an inductive qualitative approach inspired by grounded theory (Gioia *et al.*, 2013; Glaser and Strauss, 1967). Through iterative coding and constant comparison, recurring patterns were identified across organizational, strategic, and institutional dimensions, enabling an integrated understanding of how theatres navigate intelligent digital transformation as mission-driven and hybrid organizations.

2 Theoretical Background

Digital transformation is increasingly conceptualized as a multidimensional process that extends beyond the mere adoption of digital technologies, encompassing changes in how organizations create value, organize knowledge, and structure decision-making processes (Matt *et al.*, 2015; Vial, 2019; Wessel *et al.*, 2021). In this context, it is essential to distinguish between related concepts that capture different levels of change. Digitization refers to the technical conversion of analog information into digital form, while digitalization describes the application of digital technologies to improve and reorganize existing processes (Brennen and Kreiss, 2016; Saarikko *et al.*, 2020). Digital transformation, in contrast, involves the deeper integration of digital technologies and knowledge into organizational strategy, governance, and value creation, potentially reshaping organizational identity and purpose (Berman, 2012; Matt *et al.*, 2015; Wessel *et al.*, 2021).

However, recent research suggests that digital transformation does not necessarily lead to radical or disruptive change in all contexts. Instead, organizations often adopt digital technologies in incremental and selective ways, embedding them into existing structures and practices (Baiyere *et al.*, 2023; Plekhanov *et al.*, 2023). From this perspective, digital transformation is not a uniform process but one that unfolds through different dynamics. Wang et al. (2021) conceptualize this through reinforcing and differentiating mechanisms: digital technologies may either strengthen existing organizational logics and practices or enable new forms of value creation and interpretation. This distinction is particularly relevant in contexts where organizational identity and value are closely tied to historically embedded practices.

These dynamics are especially pronounced in non-profit and public organizations, where value creation extends beyond economic outcomes. Cultural institutions operate within public value frameworks that emphasize accessibility, participation, cultural preservation, and social impact (Bakhshi and Throsby, 2012). In such settings, digital technologies influence not only efficiency but also how knowledge is generated, shared, and embedded in organizational processes, as well as how relationships with stakeholders are managed (Hanisch *et al.*, 2023; Saarikko *et al.*, 2020). As a result, digital systems function as structuring elements that shape governance arrangements and organizational priorities, rather than acting as neutral technological tools.

Theatre organizations represent a distinctive case within this broader transformation. As historically embedded cultural institutions, they are strongly shaped by institutional legacies, artistic traditions, and path-dependent routines that prioritize continuity, legitimacy, and public responsibility (Bonet and Schargorodsky, 2018; Johnson, 2007; Schreyögg *et al.*, 2011). These characteristics create a structural tension between the transformative potential of digital technologies and the preservation of theatre as a live, embodied, and co-present art form (Lampel *et al.*, 2000). Consequently, the adoption and integration of digital practices are mediated by institutional logics that balance artistic integrity with organizational adaptation.

In this context, digitalization in theatres is often associated with the use of digital tools to support communication, audience engagement, and the dissemination of cultural content, while more comprehensive transformation remains limited. Digital innovation - such as streaming, hybrid formats, or digitally mediated performances - introduces new possibilities for engagement and experimentation but does not necessarily imply a fundamental reconfiguration of organizational identity or value creation (Antonopoulou and Begkos, 2020;

Jiao *et al.*, 2025). As previous studies suggest, such developments frequently remain project-based and unevenly integrated into institutional structures (Mosse *et al.*, 2022).

Building on these perspectives, this study conceptualizes digital transformation in theatres as a knowledge-based and value-oriented process shaped by institutional conditions, governance capacity, and organizational learning. Rather than assuming disruptive change, it examines how digital technologies are selectively integrated into existing practices, and how this integration contributes to cultural and social value creation within mission-driven organizational contexts.

3 Methodology

This study adopts a qualitative research design to examine how theatres interpret, govern, and strategically integrate digital technologies and knowledge practices in the context of mission-driven value creation. Rather than treating digitalization as a purely technological or efficiency-driven process, the research focuses on how digital knowledge is embedded into organizational practices, governance structures, and cultural value creation. Particular attention is given to how theatres navigate the tension between continuity and change, and how digital transformation unfolds through selective and context-dependent integration.

Given the exploratory nature of the research and the limited empirical understanding of digital transformation in mission-driven cultural organizations, the study applies an inductive approach inspired by grounded theory Glaser and Strauss (1967) and the Gioia methodology (Gioia *et al.*, 2013). This approach enables the systematic development of theoretical insights from empirical data, while maintaining transparency and rigor in the analysis. Data collection and analysis were conducted iteratively, allowing emerging themes to inform subsequent data gathering until theoretical saturation was reached.

3.1. Data collection

Data were collected through 28 semi-structured interviews conducted between December 2023 and November 2024, and in November 2025. The sample includes a diverse group of professionals from the European theatre sector, such as theatre directors, executive and financial managers, marketing and communication specialists, artists, researchers, and digital service providers. This diversity ensured a comprehensive understanding of how digitalization is interpreted and implemented across different organizational roles and institutional contexts.

Participants were selected using purposeful sampling (Palinkas *et al.*, 2015) to identify individuals with relevant experience in digital initiatives, organizational strategy, and artistic production. Attention was paid to institutional diversity, including both large, publicly funded theatres and smaller, independent or experimental organizations. This approach made it possible to capture variations in digitalization practices shaped by differences in resources, governance structures, and audience orientation.

The interviews lasted 45-60 minutes and were conducted both online and offline. All interviews were audio-recorded, transcribed verbatim, and anonymized to ensure confidentiality. An interview guide was used to structure the conversations around key themes, including organizational background, digital practices, strategic approaches, governance structures, and perceived challenges and opportunities related to digital transformation. The semi-structured format allowed flexibility while ensuring comparability across interviews.

To enhance the credibility of the findings, additional publicly available materials - such as organizational websites, reports, and digital content - were reviewed to support data triangulation.

3.2. Data analysis

The data analysis followed an inductive, iterative process based on the Gioia methodology (Gioia *et al.*, 2013). First, open coding was conducted using NVivo software, focusing on how participants describe and interpret digitalization, digital transformation, and related practices. These first-order concepts closely reflected the language and experiences of interviewees.

In the next stage, second-order themes were developed by grouping and interpreting these concepts through relevant theoretical lenses, including institutional logics, public value, and knowledge-based perspectives on digital transformation. This step enabled the identification of recurring patterns related to operational digitalization, selective integration, organizational learning, and governance.

Finally, these themes were aggregated into higher-level dimensions that structure the findings. This process revealed how digitalization in theatres operates primarily as a knowledge-based, value-oriented process embedded in existing institutional frameworks, and how transformation is shaped by organizational capacities, governance arrangements, and historically embedded practices.

Through this approach, the analysis provides an empirically grounded understanding of how theatres navigate digital transformation as mission-driven and hybrid organizations, where change emerges not through disruption but through the gradual integration of digital knowledge into existing value creation logics.

4 Results

The findings reveal how theatres interpret and implement digitalization as a knowledge-based and value-oriented process embedded in existing organizational practices. Rather than unfolding as a uniform or technology-driven transformation, digitalization appears across multiple, interconnected dimensions, ranging from operational applications to strategic integration and organizational learning. The following sections present these patterns by examining (1) the role of digitalization in operational and support activities, (2) the selective and limited nature of digital transformation, (3) the function of digital experimentation in organizational learning, (4) the importance of governance and leadership, and (5) the reinforcing dynamics that shape value-oriented digital transformation in theatres.

4.1 Digitalization as operational knowledge infrastructure

The findings indicate that theatres primarily interpret and implement digitalization within operational and support activities, including communication, ticketing, audience management, marketing, and administrative coordination. This pattern was further accelerated during the COVID-19 pandemic, when digital tools became essential for maintaining communication, coordination, and audience engagement under conditions of restricted physical access. In these areas, digital technologies function as knowledge infrastructures that structure information flows, improve transparency, and enable continuous interaction with audiences, partners, and internal stakeholders.

Rather than being driven mainly by efficiency considerations, these practices contribute to cultural and social value creation. By broadening access to performances, maintaining institutional visibility, and supporting ongoing audience relationships, digitalization reinforces theatres' public missions and their role as cultural intermediaries. Digital platforms and communication tools, for instance, extend the reach of performances and facilitate engagement beyond physical attendance, while also supporting the dissemination and preservation of cultural content.

At the same time, digitalization remains incremental and uneven. It is shaped by legacy systems, resource constraints, uneven digital capabilities, and the cyclical nature of theatrical production processes. While the pandemic intensified the use of digital tools, it did not fundamentally alter these structural conditions. As a result, digital practices are typically embedded into existing routines without significantly transforming organizational structures or value creation logics. This reflects a pattern of adaptation in which digital technologies are integrated in ways that support continuity rather than enable systemic change.

4.2. Selective integration and limited digital transformation

More advanced forms of digital transformation - understood as the strategic integration of digital knowledge into organizational priorities, governance arrangements, and value creation processes - are less prevalent across theatres. Instead of pursuing disruptive or technology-driven change, theatres tend to adopt selective and adaptive approaches that embed digital tools within established institutional frameworks.

This pattern is closely linked to the sector's value orientation. Theatres operate within a context where legitimacy, continuity, artistic integrity, and responsibility toward audiences and funders are central. As a result, digital initiatives are evaluated not only in terms of efficiency or innovation, but also in relation to their alignment with cultural missions and public value. This often limits the scope of transformation, as radical change may conflict with institutional identity or stakeholder expectations.

Digital technologies are therefore primarily used to complement and extend existing practices rather than redefine them. Even when new tools are introduced, they are typically integrated in ways that preserve established governance structures and artistic processes. Consequently, digital transformation remains partial, gradual, and strongly conditioned by institutional logics.

4.3. Digital experimentation and organizational learning

Digital experimentation plays a significant role in how theatres engage with emerging technologies and practices. Hybrid formats, streaming, and digitally mediated artistic productions create opportunities to explore new forms of audience engagement, collaboration, and cultural expression. These initiatives contribute to

organizational learning by generating knowledge about how digital tools can support both artistic and managerial processes.

However, such experimentation is typically project-based and loosely connected to core organizational structures. It often depends on individual initiatives, temporary funding, or specific productions rather than long-term strategic planning. As a result, the knowledge generated through these experiments is only partially integrated into governance, limiting its broader organizational impact.

This highlights a key finding: intelligent transformation in theatres depends less on the availability of technology and more on the ability to absorb, interpret, and strategically integrate digital knowledge. Without mechanisms to institutionalize learning, experimentation remains isolated and its transformative potential underutilized.

4.4. Governance, leadership, and value-oriented transformation

Leadership and governance capacity emerge as critical enabling conditions for bridging operational digitalization and broader transformation. Where digitalization is explicitly framed as a tool for mission fulfilment, knowledge sharing, and long-term public value creation, it is more likely to be embedded in strategic thinking and cross-functional collaboration.

In such cases, digital practices support not only operational efficiency but also organizational learning and stakeholder engagement. They become part of a broader reflection on how theatres create and sustain cultural relevance in a digitally mediated environment.

Conversely, fragmented governance structures, unclear responsibilities, and limited strategic coordination constrain the integration of digital knowledge. These conditions reinforce incremental, path-dependent practices and limit the potential for more comprehensive transformation. This suggests that governance is not only a structural issue but also a key determinant of how digitalization is interpreted and enacted within the organization.

4.5. Reinforcing dynamics and intelligent transformation

Across the findings, digitalization in theatres primarily operates through reinforcing dynamics that extend and stabilize existing value creation logics. Digital tools and practices are most consistently embedded in operational and support activities, where they enhance communication, coordination, and audience engagement without fundamentally altering core organizational structures or artistic processes.

In this context, digital knowledge systems contribute to the preservation and dissemination of cultural value by improving accessibility, supporting participation, and maintaining relationships with audiences and stakeholders. These practices strengthen institutional continuity and public mission fulfilment, rather than driving disruptive or large-scale transformation.

At the same time, the limited integration of digital initiatives into strategic governance and decision-making processes constrains the emergence of more transformative dynamics. While digital experimentation introduces new forms of engagement and generates organizational knowledge, its project-based and weakly institutionalized nature reduces its influence on long-term value creation and organizational priorities.

These patterns indicate that digitalization in theatres is predominantly aligned with existing institutional logics, where legitimacy, continuity, and responsibility toward audiences and funders shape how digital tools are interpreted and applied. As a result, digital transformation tends to unfold as a gradual and selective process, in which new practices are incorporated in ways that reinforce rather than redefine organizational identities.

The findings further show that the extent to which digitalization moves beyond reinforcing dynamics depends on organizational capacities to absorb, interpret, and strategically integrate digital knowledge. Where such capacities are limited, digital practices remain fragmented and operationally focused; where they are more developed, digitalization is more likely to be connected to broader processes of reflection, coordination, and value-oriented development.

5 Conclusion

This study examined how theatres interpret, govern, and integrate digital technologies and knowledge practices in the context of mission-driven value creation. The findings show that digitalization is primarily embedded in operational and support activities, where digital tools function as knowledge infrastructures that enhance communication, coordination, and audience engagement. These practices contribute less to efficiency alone and more to cultural and social value by improving access, sustaining relationships, and reinforcing theatres' public missions.

At the same time, digitalization remains incremental and uneven, shaped by legacy systems, limited capabilities, and the cyclical nature of theatrical production. More advanced forms of digital transformation—defined as the strategic integration of digital knowledge into governance and value creation—are less prevalent. Instead, theatres adopt selective and adaptive approaches that embed digital tools within existing institutional frameworks, reflecting an orientation toward continuity, legitimacy, and public responsibility rather than disruptive change.

Digital experimentation plays a role in organizational learning, particularly through hybrid formats and streaming, but remains largely project-based and weakly institutionalized. As a result, its impact on long-term governance and strategy is limited. Across these patterns, the findings indicate that intelligent transformation depends less on technology itself and more on organizational capacities to absorb, interpret, and strategically integrate digital knowledge.

Overall, digital transformation in theatres primarily reinforces existing value creation logics rather than redefining them. Digital knowledge systems extend access, support cultural preservation, and strengthen audience relationships, contributing to institutional resilience while maintaining the core characteristics of theatre as a live, mission-driven art form. These findings highlight that in non-profit and hybrid contexts, digital transformation unfolds as a value-oriented, knowledge-based process shaped by governance, institutional logics, and public mission.

6 Acknowledgements

The research was supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences.

7 Ethics declaration

This research was approved by Corvinus University of Budapest, Vice-Rector for Faculty and Research, Research Ethics Committee with the approval number: KRH/61/2025 dated 26.02.2025.

8 AI declaration

During the preparation of this manuscript, the author used QuillBot for paraphrasing and language refinement. This tool was used to improve the clarity, fluency, and consistency of the English text. The output generated by the tool was reviewed, critically assessed, and edited by the author before inclusion in the submitted manuscript. The author takes full responsibility for the final content, interpretation, arguments, and conclusions presented in the article.

References

Antonopoulou, K. and Begkos, C. (2020), “Strategizing for digital innovations: Value propositions for transcending market boundaries”, Technological Forecasting and Social Change, Elsevier Inc., Vol. 156, doi: 10.1016/j.techfore.2020.120042.

Baiyere, A., Grover, V., Lyytinen, K.J., Woerner, S. and Gupta, A. (2023), “Digital ‘x’—Charting a Path for Digital-Themed Research”, Information Systems Research, INFORMS Inst.for Operations Res.and the Management Sciences, Vol. 34 No. 2, doi: 10.1287/isre.2022.1186.

Bakhshi, H. and Throsby, D. (2012), “New technologies in cultural institutions: Theory, evidence and policy implications”, International Journal of Cultural Policy, Vol. 18 No. 2, pp. 205–222, doi: 10.1080/10286632.2011.587878.

Berman, S.J. (2012), “Digital transformation: opportunities to create new business models”, Strategy & Leadership, Vol. 40 No. 2, pp. 16–24, doi: 10.1108/10878571211209314.

Bonet, L. and Schargorodsky, H. (2018), Theatre Management: Models and Strategies for Cultural Venues, Vol. Rapport no. 06-2018, Kunnskapsverket, Elverum, Norway.

Brennen, J.S. and Kreiss, D. (2016), “Digitalization”, in Jensen, K.B., Craig, R.T., Pooley, J.D. and Rithenbuhler, E.W. (Eds.), The International Encyclopedia of Communication Theory and Philosophy, John Wiley & Sons, Inc., pp. 1–11, doi: https://doi.org/10.1002/9781118766804.wbiect111.

Ford, V. and Mandviwalla, M. (2020), “Can digital engagement transform the performing arts?”, Proceedings of the Annual Hawaii International Conference on System Sciences, Vol. 2020-January, IEEE Computer Society, pp. 4296–4305, doi: 10.24251/hicss.2020.526.

Gioia, D.A., Corley, K.G. and Hamilton, A.L. (2013), “Seeking Qualitative Rigor in Inductive Research: Notes on the Gioia Methodology”, Organizational Research Methods, Vol. 16 No. 1, pp. 15–31, doi: 10.1177/1094428112452151.

Glaser, B.G. and Strauss, A.L. (1967), The Discovery of Grounded Theory: Strategies for Qualitative Research, Aldine, Chicago.

Hanisch, M., Goldsby, C.M., Fabian, N.E. and Oehmichen, J. (2023), “Digital governance: A conceptual framework and research agenda”, Journal of Business Research, Vol. 162, p. 113777, doi: https://doi.org/10.1016/j.jbusres.2023.113777.

Hawthorne, K. (2023), Digital Theatre – Strategies and Business Models in European Theatre.

Jiao, H., Wang, T., Libaers, D., Yang, J. and Hu, L. (2025), “The relationship between digital technologies and innovation: A review, critique, and research agenda”, Journal of Innovation and Knowledge, Elsevier B.V., Vol. 10 No. 1, doi: 10.1016/j.jik.2024.100638.

Johnson, V. (2007), “What is organizational imprinting? Cultural entrepreneurship in the founding of the Paris opera”, American Journal of Sociology, Vol. 113 No. 1, pp. 97–127, doi: 10.1086/517899.

Lampel, J., Lant, T. and Shamsie, J. (2000), “Balancing Act: Learning from Organizing Practices in Cultural Industries”, Organization Science, Vol. 11 No. 3, pp. 263–269.

Matt, C., Hess, T. and Benlian, A. (2015), “Digital Transformation Strategies”, Business & Information Systems Engineering, Vol. 57 No. 5, pp. 339–343, doi: 10.1007/s12599-015-0401-5.

Mosse, R., Janke, J., König, J., Stein, C. and Tecklenburg, N. (2022), “Viral Theatres’ Pandemic Playbook - Documenting German theatre during COVID-19”, International Journal of Performance Arts and Digital Media, Routledge, Vol. 18 No. 1, pp. 105–127, doi: 10.1080/14794713.2022.2031800.

Plekhanov, D., Franke, H. and Netland, T.H. (2023), “Digital transformation: A review and research agenda”, European Management Journal, Elsevier Ltd, Vol. 41 No. 6, pp. 821–844, doi: 10.1016/j.emj.2022.09.007.

Saarikko, T., Westergren, U.H. and Blomquist, T. (2020), “Digital transformation: Five recommendations for the digitally conscious firm”, Business Horizons, Vol. In Press, doi: https://doi.org/10.1016/j.bushor.2020.07.005.

Schreyögg, G., Sydow, J. and Holtmann, P. (2011), “How history matters in organisations: The case of path dependence”, Management and Organizational History, Vol. 6 No. 1, pp. 81–100, doi: 10.1177/1744935910387030.

Vial, G. (2019), “Understanding digital transformation: A review and a research agenda”, The Journal of Strategic Information Systems, Vol. 28 No. 2, pp. 118–144, doi: https://doi.org/10.1016/j.jsis.2019.01.003.

De Voldere, I., Romainville, J.-F., Knotter, S., Durinck, E., Evrim, E., Le Gall, A., Kern, P., et al. (2017), Mapping the Creative Value Chains - A Study on the Economy of Culture in the Digital Age.

Wang, Y., Xiuping, S. and Zhang, Q. (2021), “Can fintech improve the efficiency of commercial banks? —An analysis based on big data”, Research in International Business and Finance, Elsevier B.V., Vol. 55 No. October 2020, p. 101338, doi: 10.1016/j.ribaf.2020.101338.

Wessel, L., Baiyere, A., Ologeanu-Taddei, R., Cha, J. and Jensen, T.B. (2021), “Unpacking the difference between digital transformation and it-enabled organizational transformation”, Journal of the Association for Information Systems, Association for Information Systems, Vol. 22 No. 1, pp. 102–129, doi: 10.17705/1jais.00655.

UX-Driven Digital Architecture for Knowledge Integration in Social-Ecological Systems

Anna Maria Crespino, Marco Lucio Sarcinella, Lorenzo Quarta
Naica Società Cooperativa, Lecce, Italy

Abstract

In the polycrisis triggered by Xylella fastidiosa in the Salento area, Italy, the regeneration of social-ecological systems depends not only on data availability, but on the ability to integrate, interpret and operationalize heterogeneous knowledge produced by multiple institutional and territorial actors. Despite the abundance of environmental data and scientific outputs, a persistent information–action gap limits the capacity of stakeholders to make coordinated and informed decisions.

This paper investigates how UX-Driven Digital Architecture can support knowledge integration and decision-making in complex socio-ecological contexts. The study adopts a qualitative case study approach combined with a research-through-design methodology, focusing on the development of the DAJS Atlas, a web-based decision-support platform for landscape regeneration.

The findings show that structuring information through actor–action–goal pathways enables users to navigate complex datasets more effectively, reduces informational fragmentation, and supports coordination among stakeholders with different objectives and knowledge frameworks. By organizing heterogeneous resources into decision-oriented flows, the platform aligns scientific, institutional, and local knowledge within shared scenarios of territorial regeneration.

The contribution of the study is twofold. Empirically, it presents a practical model for designing digital platforms that support regenerative strategies in social–ecological systems. Conceptually, it advances the understanding of User Experience as a strategic component of digital architectures, positioning UX as a systemic layer for knowledge integration, coordination, and long-term transformation in polycrisis contexts.

Keywords: Regenerative approach; Social-Ecological Systems; Decision Support System; User Experience and Research; Resilient Innovation Ecosystem.

Paper type: Practical Paper

1 Introduction

The spread of *Xylella fastidiosa* in the Salento area, located in the Apulia region (Southern Italy), has generated a condition of systemic disruption affecting ecological processes, productive structures, and territorial governance. Beyond the immediate agronomic impact, the crisis has altered the relationships between landscape management, local economies, and institutional decision-making, exposing structural fragilities in how knowledge is produced, shared, and used across stakeholders.

In this context, a significant information–action gap has emerged. Despite the availability of large volumes of environmental data, GIS-based analyses, and scientific reports, stakeholders often struggle to translate such information into coordinated and context-sensitive decisions. This gap reflects a broader challenge in knowledge systems, where the usability of information depends on its credibility, salience, and legitimacy (Cash et al., 2003).

Effective responses in social–ecological systems require the integration of heterogeneous forms of knowledge across institutional and disciplinary boundaries, particularly under conditions of uncertainty and complexity (Moallemi et al., 2023). However, this integration is often hindered by the fragmentation of data sources, the diversity of stakeholder perspectives, and the lack of mechanisms to align information with real decision-making contexts.

Digital platforms and decision-support systems are increasingly used to address complexity in social–ecological systems. However, many of these systems remain focused on data integration and visualization, while providing limited support for knowledge interpretation, coordination, and action. As a result, decision-making processes remain fragmented, particularly in multi-actor environments characterized by diverse objectives and epistemic perspectives.

This paper investigates how UX-Driven Digital Architecture can contribute to addressing this challenge by positioning User Experience (UX) as a strategic component of digital knowledge infrastructures. Rather than focusing solely on usability, this approach leverages UX principles to structure information architectures, interaction flows, and decision-making pathways that enable knowledge integration and coordination across stakeholders.

The study adopts a qualitative case study approach combined with a research-through-design methodology, focusing on the DAJS Atlas, a web-based decision-support platform developed to support landscape regeneration in the post-Xylella context. The platform integrates heterogeneous datasets related to environmental, agricultural, and socio-economic conditions, and serves as an empirical setting to explore how UX-driven design can transform data availability into actionable knowledge.

The contribution of this paper is twofold. First, it provides empirical evidence on how UX-centred design can support knowledge integration and coordinated decision-making in complex social–ecological systems. Second, it advances a conceptual interpretation of UX as a systemic design layer within digital architectures, enabling alignment between scientific, institutional, and local forms of knowledge in polycrisis contexts.

2 Background and Related Works

2.1 Knowledge Integration in Social-Ecological Systems

Social–ecological systems (SES) are complex and adaptive systems in which ecological processes, social dynamics, and institutional arrangements are deeply interconnected. Addressing challenges such as climate change, biodiversity loss, and resource management therefore requires integrative approaches capable of bridging disciplinary and sectoral boundaries. This need is emphasized by Moallemi et al. (2023), who argue that effective responses in human–natural systems depend on combining multiple forms of knowledge within decision processes under uncertainty.

However, knowledge integration in SES remains difficult. Different disciplines rely on distinct conceptual vocabularies, methods, and evaluative criteria, often producing semantic ambiguities and epistemic misalignments. A foundational contribution to this problem is Star & Griesemer (1989), who introduced the concept of boundary objects as artifacts flexible enough to adapt to local needs while robust enough to retain a common identity across communities.

Building on this perspective, the literature on sustainability and science–policy relations shows that knowledge integration is not only a matter of communication, but of designing systems that make knowledge usable across contexts. In Cash et al. (2003) argue that effective knowledge systems must balance credibility, salience, and legitimacy if they are to inform action across science, policy, and practice.

At the same time, integrating scientific, expert, and local knowledge introduces further complexity because these knowledge forms may generate divergent interpretations of the same phenomenon. Recent work on co-production stresses that such divergence should not be treated as noise to eliminate, but as a condition to be managed through iterative and collaborative processes. This is the central argument of Norström et al. (2020).

From a governance perspective, collaborative arrangements are increasingly seen as essential for enabling knowledge integration and supporting resilience in SES. Yet collaboration alone is insufficient if it does not translate into effective coordination and action. Wyborn (2015) reframes adaptive governance as a co-productive process shaped by social, contextual, and normative relations rather than by formal coordination alone.

Overall, the literature suggests that knowledge integration in social–ecological systems is not merely a technical issue of aggregating data, but a systemic challenge involving epistemic alignment, stakeholder coordination, and governance design. This creates the need for approaches that can structure heterogeneous knowledge into meaningful and actionable forms, bridging the gap between information availability and effective decision-making.

2.2 Digital Platforms as Knowledge Infrastructures

Digital platforms and decision-support systems (DSS), particularly those applied to environmental and territorial contexts, have become central tools for addressing complexity in social–ecological systems. These systems enable the integration, visualization, and analysis of heterogeneous datasets, supporting planning and decision-making processes across multiple domains. However, despite their technical potential, many environmental decision-support systems face significant limitations in terms of usability, adoption, and long-term effectiveness. As highlighted in Walling & Vaneeckhaute (2020), a large proportion of DSS initiatives are underused or abandoned, often due to a mismatch between system design and stakeholder needs, as well as insufficient support for real-world decision processes.

These limitations reflect a broader issue in the design of digital systems, where the emphasis is placed on data availability and analytical capabilities rather than on how information is structured, interpreted, and operationalized by users. In complex decision-making environments, the availability of data does not

automatically translate into actionable knowledge, particularly when users must navigate large volumes of heterogeneous information and reconcile multiple objectives.

To address this gap, recent research suggests that digital platforms should be understood not merely as tools or applications, but as knowledge infrastructures that is, socio-technical systems composed of people, technologies, and institutional arrangements that collectively enable the production, sharing, and use of knowledge. As discussed in Borgman & Groth (2025), knowledge infrastructures shape how data are interpreted, reused, and transformed into knowledge, influencing the capacity of different actors to engage with information and participate in decision-making processes.

This perspective builds on earlier work on knowledge infrastructures, which conceptualizes them as the underlying frameworks that sustain knowledge production and coordination across communities (Edwards, P. N. et al., 2013). Within this framework, digital platforms are not neutral containers of data, but active components of socio-technical systems that structure interactions, define access to information, and influence how knowledge is mobilized.

Recent studies in platform research further reinforce this view by conceptualizing digital platforms as infrastructural systems that embed social, technical, and institutional relations. In Pierson (2021), platforms are described as deeply embedded in social practices, shaping not only communication and coordination but also trust, governance, and value creation.

At the same time, research on environmental governance highlights that the effectiveness of decision-making processes depends not only on access to information, but on the quality of stakeholder interaction and participation. As shown in Newig et al. (2023), meaningful participation is associated with improved environmental outcomes, suggesting that systems supporting decision-making must enable not only access to data but also shared understanding and coordination among actors.

Despite these advances, many digital platforms still struggle to support knowledge integration and coordinated action in multi-stakeholder environments. This limitation reflects a persistent gap between data availability and its practical usability, where information is not adequately structured to support sense-making and decision-making processes.

These challenges highlight the need to move beyond a purely data-centric view of digital platforms toward an approach that emphasizes how knowledge is structured, navigated, and made meaningful for different users. In this sense, digital platforms should be conceived as environments that support sense-making, coordination, and alignment across stakeholders, laying the groundwork for design approaches that position interaction and user experience (UX) as central components of knowledge infrastructures.

2.3 From User Experience to UX-Driven Digital Architecture

UX-Driven Digital Architecture refers to the use of UX principles to structure information architectures, interaction flows, and decision-making pathways to enable knowledge integration and coordination in complex systems. This perspective extends traditional Human–Computer Interaction approaches Dix, A. (2026) by shifting the focus from interface usability to system-level orchestration of meaning, where interaction design becomes a mechanism for organizing distributed knowledge and supporting collective decision processes.

In complex socio-technical environments, usability alone is insufficient when users must interpret heterogeneous and uncertain datasets. Building on Cash et al. (2003), knowledge becomes actionable across disciplinary and institutional boundaries only when it satisfies three conditions: credibility, salience, and legitimacy. This implies that digital systems must be designed to align how information is structured and the contexts with how it is interpreted and used.

From this perspective, UX-Driven Digital Architecture can be conceptualized as comprising three interrelated components:

- Information structuring layer, which organizes heterogeneous datasets into coherent and meaningful configurations aligned with user needs and decision contexts;
- Interaction layer, which defines how users navigate, access, and interpret information through guided flows rather than isolated data exploration;
- Decision pathways layer, which connects user intentions (actor–action–goal) with relevant data, indicators, and knowledge resources, enabling coordinated and context-aware decision-making.

These components operate together to transform fragmented information environments into structured knowledge systems that support sense-making and coordination across stakeholders.

Furthermore, this approach aligns with principles of knowledge co-production Szetey et al. (2023), where different forms of knowledge are integrated through iterative and collaborative processes. In this sense, UX-Driven Digital Architecture can be interpreted as an operational translation of co-production principles into

design practice, enabling the alignment of scientific, institutional, and local knowledge within shared decision frameworks.

Rather than functioning as a layer of interface optimization, UX becomes a systemic design strategy that enables coordination across actors, datasets, and institutional boundaries, particularly in contexts characterized by high uncertainty and fragmentation.

3 Research Design and Methodological Approach

3.1 Research aims and steps

To extend the current state of the art of User Experience into the domain of social–ecological systems, this study investigates how UX-Driven Digital Architecture can support knowledge integration among multiple institutional and territorial actors operating within the polycrisis triggered by Xylella fastidiosa in Salento. The aim is to explore how UX can be repositioned as a strategic component in the design of digital systems supporting knowledge integration and decision-making in complex social–ecological contexts. Specifically, the study examines how a UX-Driven Digital Architecture can contribute to structuring heterogeneous knowledge, enabling coordination among stakeholders, and supporting decision-making processes.

To address this objective, the research adopts a qualitative case study approach combined with a research-through-design perspective. A case study is a suitable method to explore a problem and how to solve it in a real organizational setting (Creswell & Creswell, 2018). This approach allows the investigation of a real-world design process while simultaneously generating conceptual insights grounded in practice. The case study focuses on the development of a UX-Driven Digital Architecture within the DAJS project, providing an empirical context to analyse how design choices influence knowledge integration and system usability. The research is structured into the following steps:

1. An exploratory phase was conducted to understand the complexity of the context and to identify the main challenges related to knowledge integration and decision-making. This phase included the analysis of existing data sources, system requirements, and stakeholder roles within the project.
2. A design-oriented phase was carried out through qualitative data collection involving semi-structured interviews with domain experts and selected end-users. This phase aimed to elicit use situations, user needs, and perceived barriers in accessing and using available information. The qualitative data collection involved a purposive sample of domain experts and end-users, including representatives from agriculture, public administration, and technical domains. A total of 21 semi-structured interviews were conducted, each lasting approximately 60 minutes. The collected data were analysed through an iterative coding process, combining open and axial coding to identify recurring patterns in user needs, decision contexts, and information use. This process led to the identification of actor–action–goal structures that informed the design of the information architecture.
3. In the implementation phase, the UX-Driven Digital Architecture translated these insights into a structured design solution, operationalised within the DAJS Atlas, where the identified patterns were systematically embedded into the system's information architecture and interaction flows.
4. A validation phase is currently being conducted through iterative feedback sessions and usability-oriented evaluations with stakeholders. This phase aims to assess the system's capacity to support knowledge integration, improve information accessibility, and facilitate coordinated decision-making. Preliminary observations indicate that structuring information through actor–action–goal pathways enhances users' ability to navigate complex datasets, reduces the need to manually assemble fragmented information, and supports more informed and coordinated decisions across stakeholder groups. This phase is still ongoing and is focuses on collecting qualitative feedback, usage insights, and interaction patterns to inform further refinement of the platform and support its progression toward a higher Technology Readiness Level (TRL).

Through this structured process, the research combines empirical investigation and design practice to generate both practical and conceptual contributions, demonstrating how UX can function as a strategic layer within digital knowledge infrastructures.

4 Social-Ecological Context

The research is situated within the socio-ecological crisis triggered by Xylella fastidiosa in the Salento region (Southern Italy), one of the most severe plant disease outbreaks affecting Mediterranean agro-ecosystems in recent decades. The spread of the bacterium has caused the widespread destruction of olive groves, generating

profound ecological, economic, and cultural impacts across the region (*Figure 1*). Recent studies highlight the scale of this crisis, estimating significant long-term economic losses and structural transformations in olive production systems (Schneider et al., 2020). Beyond its direct agricultural impact, the Xylella outbreak has triggered a systemic disruption affecting the relationships between landscape, local identity, agricultural practices, and regional economies. This configuration reflects the characteristics of complex social–ecological systems, where environmental, social, and institutional dimensions are deeply interconnected and evolve dynamically over time. In such contexts, effective responses require the integration of heterogeneous forms of knowledge and the coordination of multiple actors operating across different scales.

The Salento context is characterized by a high degree of complexity, both in terms of environmental variables and data availability. A large volume of information is generated through environmental monitoring systems and Geographic Information Systems (GIS), covering aspects such as soil conditions, water availability, climate dynamics, ecosystem services, and agricultural suitability. Recent research emphasizes how advanced monitoring technologies contribute to the generation of high-resolution datasets for tracking the spread and impact of Xylella fastidiosa, further increasing the volume and heterogeneity of available information (Hadjichristodoulou et al., 2024). However, the availability of data does not automatically translate into actionable knowledge. Information is often distributed across multiple institutional platforms and technical systems, requiring advanced expertise to be accessed and interpreted. This fragmentation creates significant barriers for stakeholders who need to make informed decisions, particularly in contexts where time-sensitive interventions and long-term planning must coexist.

Figure 1: Impact of Xylella fastidiosa on olive landscapes in the Salento area. The widespread dieback of olive trees illustrates the ecological and socio-economic disruption caused by the outbreak. Sjor, CC BY-SA 4.0 <https://creativecommons.org/licenses/by-sa/4.0>, via Wikimedia Commons.

More specifically, three main criticalities can be identified: (i) data fragmentation across platforms and institutions limits the possibility of constructing a coherent and integrated understanding of the territory; (ii) the technical complexity of available tools restricts their usability to a limited number of expert users, excluding non-specialist stakeholders such as farmers or local decision-makers; (iii) the absence of structured decision-support mechanisms hinders coordination among actors, each operating according to different objectives, priorities, and knowledge frameworks. These challenges are further amplified by the multi-actor nature of the system. The regeneration process involves a wide range of stakeholders, including farmers, researchers, public administrations, policy makers, and technical experts. Each of these actors contributes different forms of knowledge and operates within distinct institutional and operational logics. As highlighted in recent studies on environmental governance, the effectiveness of decision-making processes depends not only on access to information, but also on the quality of stakeholder interaction and coordination (Newig et al., 2023).

In this context, existing digital platforms often function primarily as data repositories or visualization tools. While they provide access to large volumes of information, they do not necessarily support its interpretation, integration, or operationalization. As a result, decision-making processes remain fragmented, reactive, and poorly coordinated. The DAJS Atlas was developed within this context as a digital platform aimed at integrating heterogeneous datasets and supporting territorial regeneration processes. It provides access to multi-scalar information, scenario-based analyses, and thematic datasets related to environmental, agricultural, and socio-economic conditions. However, while the availability of an integrated platform represents a necessary step, it is

not sufficient to ensure effective knowledge integration. The key challenge lies in how information is structured, navigated, and aligned with the cognitive and operational needs of different users. Within this complex socio-ecological context, the UX-Driven Digital Architecture approach is introduced to transform data availability into actionable knowledge, by structuring information architectures and interaction flows that support sense-making, coordination, and shared decision-making across stakeholders.

5 Case Study: UX-Driven Digital Architecture in the DAJS Atlas platform

5.1 Context: the post-Xylella landscape

The DAJS Atlas case emerged within a context where ecological collapse was accompanied by operational and governance complexity. Following the spread of Xylella fastidiosa, large portions of the Salento olive landscape experienced severe degradation, creating urgent needs for replanting strategies, land-use reconfiguration, economic diversification, and environmental restoration. These needs generated multiple technical questions that could not be addressed through a single disciplinary lens.

Stakeholders required access to heterogeneous information sources, including soil suitability data, hydrological constraints, climatic trends, biodiversity indicators, energy infrastructure, agricultural productivity, protected areas, and regulatory constraints. These datasets existed across separate institutional repositories, often in incompatible formats and with different update cycles. While technically available, they were difficult to combine into coherent decision processes.

At the same time, actors operated with divergent priorities. Farmers needed practical guidance on crop substitution, irrigation feasibility, and economic viability. Public administrations required evidence for territorial planning, incentives, and compliance procedures. Technical professionals needed spatial diagnostics and normative references. Researchers sought exportable datasets and analytical continuity. As a result, the core challenge was not merely data scarcity, but the absence of an accessible architecture capable of translating distributed information into coordinated action.

Existing GIS environments were largely oriented toward expert use. They exposed map layers, technical metadata, and complex controls, but rarely reflected how non-specialist users formulate problems in practice. Users typically reasoned in terms such as “What can I cultivate here?”, “Which constraints apply to this parcel?”, “Where are similar successful cases?”, or “How will climate conditions evolve?”. The mismatch between system logic and user logic created a persistent information–action gap.

The post-Xylella landscape therefore represented an ideal setting to test whether UX-driven architectural principles could reorganize complex territorial knowledge into navigable decision pathways aligned with real stakeholder intentions.

5.2 UX-Driven Digital Architecture configuration

The UX-Driven Digital Architecture was implemented in DAJS Atlas as a modular web platform composed of three interconnected functional domains: (a) a scenario exploration module (DSS), (b) an interactive territorial intelligence module (WebGIS), and (c) a documentary knowledge repository.

- The first module, DSS (Decision Support Scenarios), was designed as an entry point for users who do not start from datasets but from intentions. Instead of navigating technical categories, users are guided through progressive selections based on actor profile, intended action, and expected objective. This interaction logic transforms user needs into structured queries linked to relevant territorial information (Figure 2).
- The second module, WebGIS, provides map-based exploration of the territory. It includes geospatial layers related to land use, ecosystem services, logistics, agricultural potential, climatic dynamics, and simulation outputs. However, unlike conventional GIS systems, access to layers is mediated through decision scenarios and task-oriented navigation rather than purely thematic menus (Figure 3).
- The third module, Repository, aggregates reports, scientific studies, crop sheets, regulatory documents, and practical guidelines related to sustainable regeneration. Search and filtering mechanisms allow users to retrieve documentary evidence connected to spatial contexts and decision pathways (Figure 4).

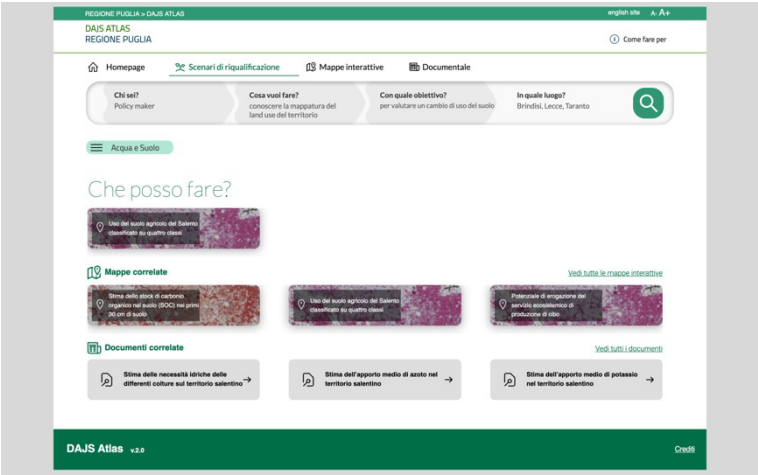


Figure 2: Component acting as a decision dashboard that guides the user in selecting the most appropriate regeneration strategies based on stated needs.

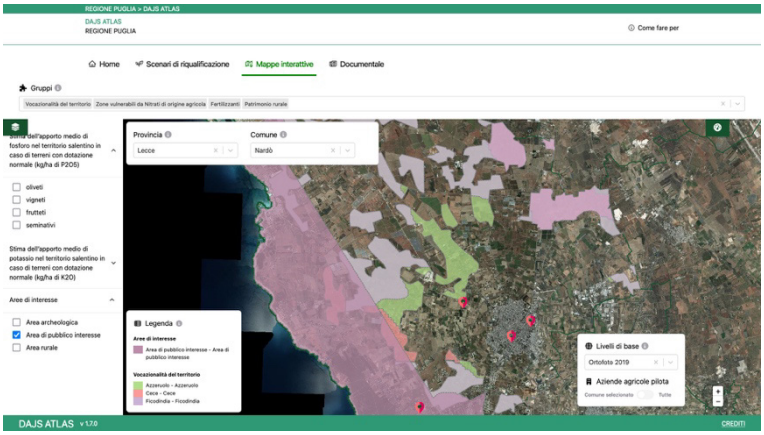


Figure 3: Interactive Maps component showing an example of the dataset of Areas of Public Interest according to the PPTR, overlaid with the dataset of potential tree and shrub species suitable for planting.

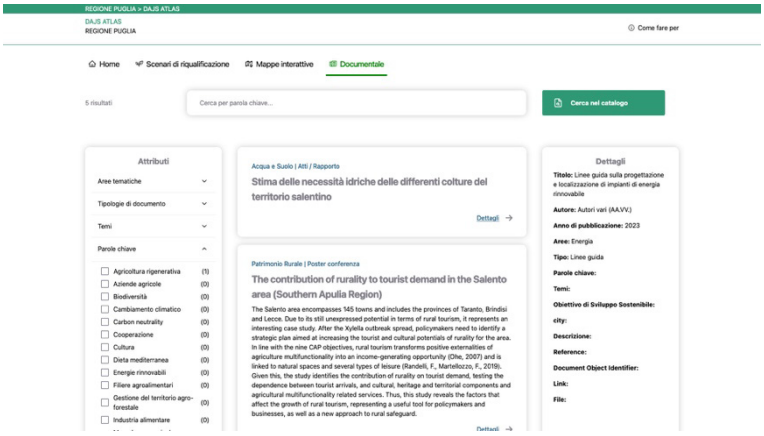


Figure 4: Repository component for searching and filtering documents related to sustainable regeneration.

From an architectural perspective, the platform was configured around dependency logic between activities. For example, geographic search enables map navigation; map navigation enables consultation of territorial constraints and case studies; case studies enable access to guidelines; map-based analysis enables forecasts and simulations; simulations enable data export for external tools. This sequencing reflects progressive knowledge acquisition rather than isolated functions.

5.3 Structuring decision-making for knowledge integration and coordination

The core design mechanism of the platform was the transformation of qualitative user research into reusable actor–action–goal pathways. Interviews and iterative design phases revealed recurring decision situations across stakeholder categories. These were converted into structured flows connecting user intentions with relevant datasets, indicators, documents, and analytical tools.

The architecture therefore combines three design principles:

- *Intentional access*: users enter through goals rather than datasets.
- *Progressive complexity*: advanced tools are revealed after contextual orientation.
- *Cross-module continuity*: maps, documents, scenarios, and analytics are linked within coherent flows.

As an example, the following connections are provided to stakeholders.

A farmer could ask which crop alternatives are viable in a damaged olive area. The system connects parcel location, pedological data, climatic suitability, irrigation availability, and cultivation guidelines.

A policy maker could evaluate where regeneration incentives should be prioritized. The platform combines land abandonment indicators, ecosystem service deficits, logistics accessibility, and socio-economic vulnerability.

A technical consultant could verify whether a land-use transition is feasible. The system links administrative boundaries, environmental constraints, normative references, and comparable case studies.

A researcher could obtain simulation outputs for external modelling. The platform enables export of forecast and KPI data generated within the system.

These pathways reduce the need for users to manually assemble fragmented information across institutions. They also create shared reference frames among actors whose objectives differ but intersect territorially. For instance, the economic resilience sought by farms, the ecological restoration sought by environmental actors, and the planning rationality sought by administrations can be examined through common geographies and evidence bases.

In this sense, UX design functions here as a translation layer between heterogeneous epistemic communities.

6 Discussion

The case extends conventional interpretations of UX by positioning it as an infrastructural capability rather than a front-end optimization practice. In complex social–ecological systems, usability alone is insufficient because the central challenge is often epistemic fragmentation: data exist but remain disconnected from decision contexts.

The DAJS Atlas suggests that UX can contribute to three higher-order functions.

First, UX can structure knowledge through pathways based on actor goals, allowing the system to filter what is relevant at specific moments.

Second, UX can support mediation among different actors interpreting the same territory through different lenses. A shared architecture enables the coexistence of these perspectives without collapsing them into a single logic.

Third, UX can enhance coordination capacity. By linking actions, data, and objectives across modules, interaction design contributes to governance performance.

This supports the broader proposition that digital platforms in sustainability transitions should be analysed not only as technologies, but as socio-technical architectures where UX plays a strategic role in aligning cognition, institutions, and action.

The DAJS Atlas case provides several practical implications for territorial governance, policy design, and digital platform development in polycrisis contexts.

For territorial systems, it demonstrates how digital infrastructures can operationalize regenerative strategies by translating fragmented datasets into decision pathways aligned with local priorities. The actor–action–goal framework offers a method to transform spatial data into actionable insights that support context-specific planning, crop diversification, and adaptive land-use transitions. This approach helps municipalities and local consortia move from reactive crisis management toward anticipatory and regenerative governance.

For decision-makers, the platform highlights the need to redesign governance mechanisms around information accessibility and interpretability. By aligning scientific outputs with institutional workflows, DAJS Atlas supports evidence-based decisions, facilitates interdepartmental coordination, and strengthens accountability across multi-level governance structures. It suggests that post-crisis regeneration policies should invest not only in data collection but also in mediation capacities that enable shared sense-making and cross-sectoral learning.

For digital platform design, the case illustrates the strategic value of UX as a governance instrument rather than a mere usability enhancement. Embedding participatory inquiry and qualitative coding into the design

process enables platforms to reflect the real cognitive and operational needs of diverse stakeholders. This reframes digital tools as dynamic knowledge infrastructures that mediate between scientific, administrative, and local forms of knowledge.

7 Conclusions

This paper examined how UX-Driven Digital Architecture can support stakeholder engagement and knowledge integration in the post-Xylella regeneration of Salento. The DAJS Atlas case shows that the main challenge in polycrisis environments is often not lack of data, but the inability to convert dispersed information into coordinated decisions.

By organizing heterogeneous resources through actor–action–goal pathways, the platform transformed GIS layers, reports, simulations, and regulatory content into navigable decision structures tailored to multiple stakeholder groups. In doing so, UX operated as a strategic layer for territorial coordination, reducing fragmentation and supporting regenerative planning.

The contribution is twofold. Empirically, the study documents a practical model for designing decision-support platforms in stressed territorial systems. Conceptually, it advances the view of UX as an infrastructural practice for aligning knowledge, institutions, and long-term transformation goals.

Future research may evaluate longitudinal impacts on adoption, inter-organizational coordination, and measurable regeneration outcomes, as well as test transferability of the model in other climate, biodiversity, or agricultural crisis contexts.

8 Acknowledgements

This contribution is the result of a collective reflection by the authors on the outcomes of the Research and Development Project “Sustainable Regeneration of Agriculture in Territories Affected by Xylella fastidiosa”, promoted by the Jonico-Salentino Agro-Food Quality District (DAJS) and funded by the Italian Ministry of Agriculture, Food Sovereignty and Forestry (MASAF).

9 Ethics Declaration

The author(s) declare that the research presented in this paper did not involve human participants, animal subjects, or the collection of sensitive personal data. Consequently, ethical clearance from an institutional review board or ethics committee was not required for this study.

10 AI Declaration

The authors used generative AI tools in a limited capacity for language refinement only. No AI tools were used for idea generation, data analysis, or interpretation. The authors fully reviewed and take responsibility for all content.

References

Borgman, C. L., & Groth, P. (2025). From Data Creator to Data Reuser: Distance Matters. Harvard Data Science Review, 7(2). https://doi.org/10.1162/99608f92.35d32cfc

Cash, D. W., Clark, W. C., Alcock, F., Dickson, N. M., Eckley, N., Guston, D. H., Jäger, J., & Mitchell, R. B. (2003). Knowledge systems for sustainable development. Proceedings of the National Academy of Sciences, 100(14), 8086–8091. https://doi.org/10.1073/pnas.1231332100

Creswell, J. W., & Creswell, J. D. (2018). Research design: Qualitative, quantitative, and mixed methods approaches (Fifth edition). SAGE.

Dix, A. (2026). Human-Computer Interaction (1st Edition).

Edwards, P. N., Jackson, S. J., Chalmers, M. K., Bowker, Bowker, G. C, Borgman, C. L., Ribes, D., & Calvert, S. (2013). Knowledge infrastructures: Intellectual frameworks and research challenges.

Hadjichristodoulou, M., Papoutsa, C., Kanetis, L., Stavrinides, M., Eliades, M., & Hadjimitsis, D. (2024). An insight into remote sensing solutions for monitoring the Xylella fastidiosa bacterium that is threatening olive trees: A literature review. In S. C. Michaelides, D. G. Hadjimitsis, C. Danezis, N. Kyriakides, A. Christofe, K. Themistocleous, & G. Schreier (A c. Di), Tenth International Conference on Remote Sensing and Geoinformation of the Environment (RSCy2024) (p. 36). SPIE. https://doi.org/10.1117/12.3037289

Moallemi, E. A., Zare, F., Hebinck, A., Szetey, K., Molina-Perez, E., Zyngier, R. L., Hadjikakou, M., Kwakkel, J., Haasnoot, M., Miller, K. K., Groves, D. G., Leith, P., & Bryan, B. A. (2023). Knowledge co-production for decision-making in human-natural systems under uncertainty. Global Environmental Change, 82, 102727. https://doi.org/10.1016/j.gloenvcha.2023.102727

Newig, J., Jager, N. W., Challies, E., & Kochskämper, E. (2023). Does stakeholder participation improve environmental governance? Evidence from a meta-analysis of 305 case studies. Global Environmental Change: Human and Policy Dimensions, 82, 102705. https://doi.org/10.1016/j.gloenvcha.2023.102705

Norström, A. V., Cvitanovic, C., Löf, M. F., West, S., Wyborn, C., Balvanera, P., Bednarek, A. T., Bennett, E. M., Biggs, R., De Bremond, A., Campbell, B. M., Canadell, J. G., Carpenter, S. R., Folke, C., Fulton, E. A., Gaffney, O., Gelcich, S., Jouffray, J.-B., Leach, M., … Österblom, H. (2020). Principles for knowledge co-production in sustainability research. Nature Sustainability, 3(3), 182–190. https://doi.org/10.1038/s41893-019-0448-2

Pierson, J. (2021). Digital platforms as entangled infrastructures: Addressing public values and trust in messaging apps. European Journal of Communication, 36(4), 349–361. https://doi.org/10.1177/02673231211028374

Schneider, K., Van Der Werf, W., Cendoya, M., Mourits, M., Navas-Cortés, J. A., Vicent, A., & Oude Lansink, A. (2020). Impact of Xylella fastidiosa subspecies pauca in European olives. Proceedings of the National Academy of Sciences, 117(17), 9250–9259. https://doi.org/10.1073/pnas.1912206117

Star, S. L., & Griesemer, J. R. (1989). Institutional Ecology, `Translations' and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. Social Studies of Science, 19(3), 387–420. https://doi.org/10.1177/030631289019003001

Szetey, K., Moallemi, E. A., & Bryan, B. A. (2023). Knowledge Co-Production Reveals Nuanced Societal Dynamics and Sectoral Connections in Mapping Sustainable Human–Natural Systems. Earth's Future, 11(9), e2022EF003326. https://doi.org/10.1029/2022EF003326

Walling, E., & Vaneeckhaute, C. (2020). Developing successful environmental decision support systems: Challenges and best practices. Journal of Environmental Management, 264, 110513. https://doi.org/10.1016/j.jenvman.2020.110513

Wyborn, C. (2015). Co-productive governance: A relational framework for adaptive governance. Global Environmental Change, 30, 56–67. https://doi.org/10.1016/j.gloenvcha.2014.10.009

A Multidimensional KPI-Based Framework for Evaluating AI-Driven Cybersecurity Solutions

Mariangela Lazoi, Marianna Lezzi, Nicolò Gianmauro Totaro
University of Salento, Lecce, Italy

Abstract

Artificial Intelligence (AI)-powered cybersecurity is increasingly adopted in complex and data-intensive environments, particularly in Internet of Things (IoT) settings, where traditional rule-based approaches struggle to cope with dynamic and unforeseen threats. While AI-based solutions demonstrate high detection accuracy, responsiveness, and operational efficiency, their evaluation remains dominated by technical performance indicators such as precision, recall, false-positive rates, and computational efficiency. This focus provides a partial view of the suitability of AI-based cybersecurity solutions, ignoring relevant issues related to energy consumption, carbon footprint, organizational impact, accountability, and regulatory compliance. To addresses this limitation, the paper adopts a conceptual and design-oriented research approach to propose a multidimensional key performance indicators (KPI)-based framework for comparing AI-driven cybersecurity solutions, considering technical, environmental, social, economic, and governance dimensions. The framework operationalises evaluation through a structured pipeline that combines three main elements: the AI-based cybersecurity solutions under evaluation, a baseline set of KPI, and expert judgements on the relative importance of the five evaluation dimensions. The framework operationalises evaluation through a structured pipeline. Heterogeneous KPIs are classified according to their direction of preference and normalised onto a common scale to ensure comparability. Normalised indicators are aggregated within each dimension to generate dimensional performance scores. Expert judgements are processed to derive dimension weights reflecting contextual priorities. Dimensional scores and weights are combined through weighted aggregation to produce both a composite sustainability-performance index and a multidimensional performance profile for each solution. The contribution of the paper lies in moving beyond accuracy-centric benchmarking by providing a transparent and replicable evaluation logic that makes trade-offs across dimensions explicit. Rather than treating sustainability as an external factor or an additional descriptive element, the framework integrates it directly into the evaluation process, enabling more balanced and context-sensitive comparison among AI-driven cybersecurity solutions. The proposed framework provides a methodological foundation for future empirical applications and supports more informed decision-making in the selection and governance of AI-based cybersecurity technologies.

Keywords: Artificial Intelligence; Cybersecurity; Sustainability; Key Performance Indicators; Multi-Criteria Decision-Making

Paper type: Academic research paper

1 Introduction

The increasing digitalisation of organisational processes has intensified exposure to cyber threats, particularly in complex environments such as those based on the Internet of Things (IoT) (Mohamed, 2025; Uddin et al., 2025). In response to this growing complexity, traditional cybersecurity approaches are increasingly complemented or replaced by Artificial Intelligence (AI)-powered solutions that rely on machine learning and deep learning techniques to detect anomalies and adapt to evolving attack patterns (Alansary et al., 2025; Mohamed, 2025). AI-based cybersecurity solutions can significantly outperform traditional approaches in terms of detection accuracy, with reported values exceeding 99% in controlled intrusion detection benchmarks (Shan, 2025; Yang et al., 2024). These performance gains are often accompanied by reductions in false-positive rates and analyst workload, contributing to more efficient Security Operations Centre (SOC) operations (BustamanteV et al., 2025; Shan, 2025). However, detection performance may degrade in heterogeneous and real-world IoT environments, where data diversity and evolving attack behaviours increase system complexity (Rafique et al., 2024). This variability highlights the limitations of evaluating AI-driven cybersecurity solutions solely through benchmark-based technical metrics (Mohamed, 2025; Rawat, 2025).

Sustainability has emerged as a key dimension of digital transformation strategies, extending beyond environmental efficiency to encompass social, economic, and governance considerations (Lezzi et al., 2025; Yu et al., 2025). However, current practices for evaluating technological solutions remain largely focused on technical performance indicators such as accuracy, precision, recall, false-positive rates, and computational efficiency (Mohamed, 2025; Rawat, 2025). Although these metrics are essential for assessing detection effectiveness, they are insufficient for revealing the impacts associated with AI-based security solutions, especially as model complexity and computational demands increase (Agarwal and Gupta, 2025; De Vries, 2023).

Concerns regarding energy consumption, carbon footprint, operational costs, transparency, and governance are still rarely integrated into cybersecurity evaluation frameworks (Achuthan et al., 2025; De Vries, 2023). Accuracy-centric assessment approaches risk overlooking environmental, social, and organisational implications that are critical for long-term sustainability (Lezzi et al., 2025; Priebe et al., 2023).

AI-driven solutions are increasingly recognised as both enablers of sustainable digital practices and sources of new sustainability challenges (Achuthan et al., 2025; Bhimavarapu, 2025). On the one hand, AI-based automation can enhance resource efficiency and organisational resilience; on the other hand, the growing computational footprint of advanced models raises concerns related to energy use, electronic waste, and unequal access to advanced infrastructures (Brevini, 2020; Gómez et al., 2023; Le Goff, 2025). Since current approaches to evaluating AI-driven cybersecurity rarely integrate these sustainability dimensions in a systematic and operational manner, the literature offers fragmented assessment practices (Henao et al., 2025; Rawat, 2025; Usman et al., 2025).

To fill this gap, this study proposes a conceptual framework based on key performance indicators (KPIs) for the multidimensional comparison of AI-driven cybersecurity solutions that integrates technical performance with environmental, social, economic, and governance considerations. The framework aims to support more balanced, transparent, and sustainability-aware evaluations of AI-based cybersecurity.

2 State-of-the-Art

Recent research on AI-driven cybersecurity has mainly examined how machine learning and deep learning techniques can improve the detection, classification, and prediction of cyber threats in increasingly complex digital ecosystems. Applications include intrusion detection, anomaly detection, malware classification, and threat intelligence, where AI models are used to identify non-linear patterns and evolving behaviours that are difficult to capture through static rule-based systems (Alansary et al., 2025; Hamid and Rahman, 2025; Uddin et al., 2025). Within this stream of research, IoT and industrial environments represent particularly relevant application contexts, due to their high connectivity, data heterogeneity, and exposure to dynamic attack surfaces (Mohamed, 2025; Uddin et al., 2025).

Evaluation practices in this literature remain primarily centred on technical effectiveness. Most studies assess AI-based cybersecurity solutions through indicators such as accuracy, precision, recall, false-positive and false-negative rates, latency, and computational efficiency (Mohamed, 2025; Rawat, 2025). These metrics are essential for demonstrating detection capability and operational feasibility, and several experimental studies report very high performance levels in benchmark-based intrusion detection scenarios (Shan, 2025; Yang et al., 2024). However, recent contributions also show that such results are often context-dependent, as model performance can be affected by data heterogeneity, deployment conditions, and evolving attack behaviours in real-world environments (Rafique et al., 2024). For this reason, optimisation-driven and hybrid approaches increasingly attempt to balance detection accuracy with latency, resource consumption, and deployment constraints in operational settings such as SOCs and industrial control systems (Ranpara et al., 2025).

While technical performance remains the dominant evaluation focus, recent studies increasingly acknowledge the relevance of sustainability considerations in AI-driven cybersecurity (Achuthan et al., 2025; Lezzi et al., 2025). The study by (Lezzi et al., 2025) emphasizes the distinction between “green cybersecurity”, which focuses primarily on environmental efficiency, and the broader notion of “sustainable cybersecurity”, which adopts a multidimensional perspective integrating environmental, social, economic, and governance dimensions. Green cybersecurity approaches concentrate on reducing energy consumption and carbon footprint through lightweight models and resource-aware optimisation, often without extending evaluation beyond technical and environmental indicators (Bhimavarapu, 2025; Rawat, 2025).

From an environmental perspective, increasingly complex AI models can entail significant energy consumption and associated carbon emissions, particularly during training and retraining cycles (Brevini, 2020; De Vries, 2023). In response, green AI approaches propose techniques such as model compression and adaptive learning to mitigate environmental impact while maintaining acceptable performance levels (Gupta et al., 2025; Rawat, 2025). Nevertheless, these approaches generally remain focused on technical optimisation, without explicitly connecting environmental indicators to broader organisational or societal evaluation frameworks.

The social dimension of AI-driven cybersecurity is comparatively underexplored and is typically addressed at a conceptual rather than operational level (Achuthan et al., 2025; Henao et al., 2025). Issues such as privacy protection, trust in automated decisions, workforce implications, and risks of bias, are discussed but rarely translate these concerns into measurable indicators suitable for comparative evaluation (Arpaci, 2024; Dr. Varsha, 2023). As a result, social considerations remain marginal in evaluation practices despite their relevance for responsible AI adoption.

Similarly, governance aspects, including accountability, transparency, explainability, robustness, and regulatory compliance, are increasingly recognised as critical for sustainable AI-driven cybersecurity deployment (BustamanteV et al., 2025; Usman et al., 2025). However, governance principles are seldom operationalised through explicit evaluation metrics or integrated with technical performance indicators, limiting their role in organisational decision-making (Königstorfer and Thalmann, 2022; Mohamed, 2025).

From an economic standpoint, several studies address trade-offs between detection performance, computational cost, and operational efficiency, particularly in relation to total cost of ownership and resource allocation in cybersecurity operations (Azrour et al., 2025; Ranpara et al., 2025). Nevertheless, economic considerations are often treated implicitly or in isolation, rather than being systematically integrated into multidimensional evaluation frameworks.

Overall, the literature reveals a fragmented evaluation landscape in which technical performance is extensively measured, while environmental, social, economic, and governance dimensions are addressed inconsistently and rarely integrated within a unified assessment logic (Achuthan et al., 2025; BustamanteV et al., 2025; Rawat, 2025). This fragmentation limits the ability of organisations to systematically assess trade-offs between AI effectiveness and sustainability-related impacts, reinforcing the need for a multidimensional evaluation framework.

3 Research Approach

This study adopts a conceptual and design-oriented research approach (Gregor and Hevner, 2013; Hevner et al., 2004) to develop a structured framework for evaluating AI-powered cybersecurity solutions from a multidimensional sustainability perspective.

Specifically, the approach is grounded in research on AI-driven cybersecurity evaluation and sustainable digital systems, which highlights the limits of accuracy-centric assessments and the need to account for environmental, social, economic, and governance considerations (Lezzi et al., 2025; Mohamed, 2025). Accordingly, the framework is derived through a twofold process. First, technical KPIs are identified from the AI-driven cybersecurity evaluation literature, where performance assessment is commonly based on indicators such as detection accuracy, false positives, and computational efficiency (Mohamed, 2025). Second, the environmental, social, economic, and governance dimensions are developed as author-defined KPI categories, conceptually informed by recurrent concerns discussed in recent literature on sustainable cybersecurity, responsible AI, and digital governance. Rather than being directly adopted from existing frameworks, these KPIs are formulated and adapted to the specific evaluation context of AI-driven cybersecurity solutions. This process allows the baseline KPI architecture to translate heterogeneous literature-based concerns into a parsimonious set of operational indicators suitable for cross-dimensional comparison of AI-based cybersecurity solutions.

The aim is to design an evaluation artefact that supports organisational decision-making by integrating heterogeneous criteria within a coherent structure (Belton and Stewart, 2002; Gregor and Hevner, 2013). To do this framework draws on multi-criteria evaluation and decision-support research for comparing alternatives characterised by multiple, potentially conflicting objectives (Belton and Stewart, 2002; Keeney and Raiffa, 1993). The scope is intentionally delimited. The framework supports transparent, context-aware evaluation by making trade-offs between technical performance and sustainability dimensions explicit (Montibeller and Von Winterfeldt, 2015; OECD et al., 2008). The framework provides a conceptual foundation that can be instantiated and assessed in future work, for example through case studies or expert-based evaluations (Belton and Stewart, 2002; Lezzi et al., 2025).

The baseline KPI set adopted in this study is conceived as a reference configuration derived from the convergence between two literature streams: technical evaluation of AI-driven cybersecurity and sustainability-oriented assessment of digital systems. The technical KPIs are directly derived from the cybersecurity and AI literature, whereas the remaining KPIs are inspired by the sustainability, responsible AI, and AI governance literature and subsequently adapted to the specific evaluative context of AI-enabled cybersecurity solutions. The selection of fifteen KPIs, equally distributed across five dimensions, reflects a parsimonious design choice aimed at ensuring structural symmetry, interpretability, and comparability across alternatives. This choice is consistent with composite-indicator construction guidelines, which emphasise the need for a clear theoretical structure, analytically meaningful indicator selection, and transparent comparison across units of analysis (OECD et al., 2008). The KPI set is not intended as an exhaustive list, but as a balanced operational baseline that captures the minimum evaluative architecture required to jointly assess technical performance and sustainability impacts in AI-enabled cybersecurity solutions. In line with the scope of the framework, the selected indicators focus primarily on the operational use phase of AI systems, combining technical, environmental, social, economic, and governance measures that are both organisationally relevant and operationally measurable. However, the

framework remains modular: depending on specific evaluation needs, users may add, substitute, or refine indicators, including more explicitly AI-specific or lifecycle-oriented KPIs related to phases such as training, fine-tuning, retraining, or dataset curation, while preserving the underlying evaluation logic of the framework.

4 A Multidimensional Framework for Evaluating AI-Powered Cybersecurity

4.1 Input layer: evaluation scope, KPIs, and expert knowledge

The proposed framework is designed to support the systematic comparison of AI-powered cybersecurity solutions through a structured, KPI-driven architecture. As illustrated in Figure 1, the framework explicitly defines its evaluation scope by identifying the elements that enter the assessment process and by structuring them into a coherent input layer. This design choice ensures that the evaluation process is transparent from the outset and that all subsequent analytical steps can be traced back to clearly specified inputs.

The primary unit of analysis within the framework is an AI-based cybersecurity solution to be evaluated, such as an intrusion detection system, anomaly detection model, or integrated cybersecurity platform deployed within an organisational or infrastructural context. Each solution is assessed within a common evaluation space defined by a predefined baseline set of KPIs.

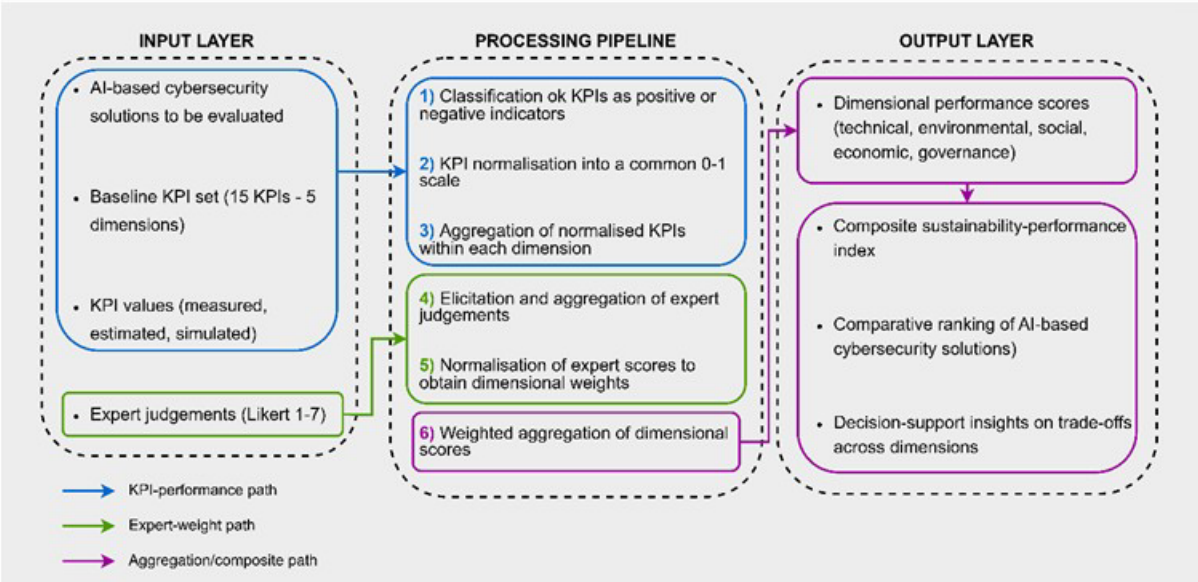


Figure 1: Framework design

The baseline KPI set is structured along five evaluation dimensions: technical, environmental, social, economic, and governance. This structure ensures consistency and comparability across all solutions under evaluation, in line with methodological recommendations for constructing composite indicators based on clearly defined dimensions and indicators (OECD et al., 2008). The framework requires that KPI values may be empirically measured, analytically estimated, or simulated, depending on data availability and evaluation context. This flexibility reflects the practical constraints of cybersecurity evaluation, where access to complete and homogeneous performance data is often limited and security metrics are typically context-dependent (Pendleton et al., 2017).

For the technical dimension, the framework considers *detection accuracy*, namely the ability of the solution to correctly identify malicious activities; *false positive rate*, namely the frequency with which legitimate events are classified as threats; and *computational efficiency*, namely the ability to maintain performance with limited computational resources (Mohamed, 2025). For the environmental dimension, the selected KPIs are *inference energy intensity*, referring to the energy consumed by the AI component during operation; *energy-to-performance ratio*, referring to the extent to which the AI component achieves detection performance relative to the energy consumed during operation; and *operational carbon intensity*, referring to the emissions associated with AI inference and model-serving activities. For the social dimension, the framework includes *sensitive-data minimisation*, namely the extent to which the solution limits the use, retention, or exposure of sensitive data in operational processing; *explanation usefulness for analysts*, namely the degree to which system outputs provide understandable and actionable support to human operators; and *human override and feedback*

support, namely the extent to which operators can supervise, contest, and intervene in system behaviour. For the economic dimension, the selected KPIs are AI *operating cost intensity*, referring to the cost attributable to the AI component during operation; *analyst productivity gain*, namely the extent to which the solution reduces triage effort and investigation time; and *cost per validated alert*, namely the average operational cost associated with generating and processing a validated security alert. Finally, for the governance dimension, the framework considers *model lineage and traceability*, namely the extent to which model versions, thresholds, data dependencies, and decision logs are documented and traceable; *drift monitoring and control readiness*, namely the ability to detect, monitor, and respond to model degradation over time; and *AI compliance and audit readiness*, namely the degree to which the solution supports auditable evidence for cybersecurity and AI-related governance requirements.

The baseline KPI set should therefore be interpreted as a general-purpose and literature-grounded reference configuration. Its role is to provide a stable and comparable evaluative structure across heterogeneous AI-driven cybersecurity solutions, while remaining sufficiently flexible to accommodate context-specific refinements. This is particularly relevant in the AI domain, where some applications may require the integration of lifecycle-specific indicators related to training, validation, deployment, monitoring, drift management, or retraining.

In addition to KPI values, the framework explicitly incorporates expert knowledge as a distinct and formalised input. Experts are asked to express judgements on the relative importance of the evaluation dimensions using a Likert scale ranging from 1 to 7, preserving interpretability and discriminative capacity in evaluative judgement tasks (Dawes, 2008; Finstad, 2010). These expert responses capture contextual priorities, such as organisational risk tolerance, regulatory sensitivity, or strategic emphasis on sustainability objectives, that cannot be inferred directly from KPI values alone.

4.2 KPI processing and dimensional scoring

The framework applies a structured KPI processing pipeline aimed at transforming heterogeneous performance indicators into analytically consistent and comparable scores. This processing phase represents the core analytical component of the framework, where raw inputs are systematically converted into dimensional performance assessments. Within this logic, the framework distinguishes between a stable evaluation architecture and context-specific implementation choices. The five dimensions and the baseline KPI structure constitute the stable backbone of the framework, while the operational instantiation may vary depending on data availability, sectoral requirements, and organisational priorities.

The first step of KPI processing consists of classifying each indicator as either a positive or a negative KPI, depending on whether higher values correspond to better or worse performance. This classification is a necessary precondition for ensuring interpretability during subsequent transformations and avoids ambiguities when aggregating indicators with opposite performance orientations, such as detection accuracy and energy consumption, as required in composite-indicator construction processes (OECD et al., 2008).

Once classified, KPI values are normalised onto a common 0-1 scale (step 2). Normalisation serves to eliminate scale effects across indicators that may originally be expressed in different units, magnitudes, or measurement ranges. By mapping all KPIs onto a shared scale, the framework ensures that no indicator dominates the evaluation solely because of its numerical properties. Importantly, the framework does not prescribe a specific mathematical normalisation formula, allowing evaluators to adopt approaches consistent with data availability and contextual constraints, provided that the directionality and relative meaning of each KPI are preserved.

Normalised KPIs are then aggregated within their respective evaluation dimensions to obtain dimensional performance scores (step 3). This intra-dimensional aggregation captures trade-offs among KPIs belonging to the same dimension. At this stage, each dimension is represented by a single score that summarises its internal performance profile, while remaining traceable to the underlying KPIs.

In parallel to KPI aggregation, expert judgements collected at the input stage are processed to derive dimension weights (step 4). Individual expert responses on dimensional importance are first aggregated and then normalised to ensure comparability across respondents. This process yields a set of weights that reflect the relative importance assigned to each dimension within a specific evaluation context, making the framework sensitive to organisational priorities without embedding fixed or universal assumptions.

A key design feature of this processing step is the explicit separation between performance scoring and importance weighting (step 5). KPI-based dimensional scores are computed independently of expert preferences, ensuring that performance evidence is not distorted by value judgements during early stages of analysis.

Through this structured KPI processing and dimensional scoring logic, the framework establishes a rigorous and traceable bridge between heterogeneous input data and higher-level evaluation outputs, preparing the ground for cross-dimensional aggregation and comparative assessment.

Once dimensional performance scores and corresponding weights have been derived, the framework proceeds to cross-dimensional aggregation (step 6). This step combines empirically grounded dimensional scores with context-specific importance weights to generate integrated evaluation outputs for each AI-based cybersecurity solution.

The first outcome of this step is the calculation of a composite sustainability-performance index, obtained through weighted aggregation of dimensional scores. Each dimension contributes to the overall index proportionally to its normalised importance weight. This structure ensures that the final composite value reflects both objective performance evidence (through KPI-derived dimensional scores) and contextual priorities (through expert-defined weights).

The framework does not collapse evaluation results into a single opaque metric. Alongside the composite index, it preserves dimensional performance scores, allowing evaluators to compare solutions at multiple analytical levels and distinguish between alternatives that achieve similar overall scores through different performance configurations. This dual-output structure supports scenario analysis and sensitivity checks, since changes in dimensional weights can show how alternative strategic priorities reshape comparative outcomes.

5 Discussion and Conclusions

This study contributes to AI-powered cybersecurity evaluation by addressing the predominance of accuracy-centric and technically focused assessment approaches. By introducing a structured, KPI-based framework that integrates technical, environmental, social, economic, and governance considerations, the proposed approach advances evaluation from isolated benchmarking towards a multidimensional and sustainability-aware perspective.

The framework enables comparative ranking and benchmarking across alternative AI-based cybersecurity solutions. Since all KPI values are normalised and aggregated according to a consistent procedure, resulting scores are directly comparable within the same evaluation exercise. This supports structured benchmarking across different models, architectures, or deployment configurations under identical evaluative assumptions, useful to provide explicit decision-support insights on trade-offs across dimensions.

The aggregation logic maintains full traceability from composite index back to dimensional scores and ultimately to underlying KPIs. This traceability strengthens interpretability and supports auditability, particularly in regulated or high-stakes cybersecurity contexts. The framework produces not only a ranking outcome, but a structured performance profile that supports informed, transparent, and reproducible evaluation of AI-based cybersecurity solutions.

From a theoretical standpoint, the framework contributes to the literature by operationalising sustainability in AI-driven cybersecurity evaluation without reducing it to a qualitative checklist or abstract principle. While prior research has highlighted sustainability-related concerns in AI-driven cybersecurity, these dimensions are often addressed in a fragmented or dimension-specific manner (Lezzi et al., 2025).The proposed framework embeds them within a coherent evaluation architecture that explicitly links KPIs, dimensional scores, expert judgement, and aggregation logic. This integration represents a methodological contribution that bridges research on AI performance evaluation, sustainable digital systems, and multi-criteria decision support, while remaining grounded in the specific requirements of cybersecurity contexts.

A key strength of the framework lies in its explicit separation between performance evidence and value-based judgement. By computing KPI-based dimensional scores independently from expert-defined importance weights, the framework avoids implicit or opaque weighting schemes that often characterise composite indicators. As a result, the framework supports reasoned deliberation rather than automated optimisation, which is particularly appropriate in cybersecurity settings characterised by uncertainty, high stakes, and regulatory constraints. This facilitates scenario analysis and sensitivity exploration, allowing organisations to examine how changes in strategic priorities, such as increased emphasis on environmental sustainability or governance robustness, alter comparative outcomes.

The framework also enables a more nuanced understanding of trade-offs inherent in AI-driven cybersecurity adoption. The dual output structure, namely dimensional profiles alongside a composite index, makes trade-offs explicit and analysable. This feature is especially relevant for organisations facing competing objectives, such as maximising detection effectiveness while minimising environmental impact or ensuring compliance with governance requirements. By preserving traceability from composite scores back to KPIs, the framework facilitates auditability and supports accountability in evaluation processes.

The framework exhibits some limitations. The framework is conceptual, and it has not yet been empirically validated through real-world case studies or comparative deployments. These limitations reflect deliberate design decisions aimed at preserving flexibility and generality, but they also point to clear directions for future research. In particular, the current KPI configuration should be understood as a transferable baseline rather than as a definitive taxonomy of AI-specific indicators. Future work may therefore extend the framework through empirical validation in different cybersecurity contexts, refinement of KPI sets for specific application domains, and the incorporation of more explicitly AI-specific or lifecycle-oriented indicators, for example related to training, validation, deployment, monitoring, and retraining. Such extensions would further strengthen the practical applicability of the framework while preserving its core conceptual structure.

Ethics declaration

This study does not involve human participants, personal data, or sensitive information. Ethical approval was therefore not required.

AI declaration

AI tools were used to support language editing and clarity of expression. The intellectual content, structure, and scientific arguments were entirely developed by the authors.

References

Achuthan, K., Sankaran, S., Roy, S., Raman, R., 2025. Integrating sustainability into cybersecurity: insights from machine learning based topic modeling. Discov Sustain 6, 44. https://doi.org/10.1007/s43621-024-00754-w

Agarwal, P., Gupta, N., 2025. AI-Driven Digital Transformation in Small and Medium-Sized Enterprises (SMEs): Resilience, Efficiency, and Growth in Industry 5.0, in: Mehra, P., Abioye, O., Kumar, P. (Eds.), Advances in Computational Intelligence and Robotics. IGI Global Scientific Publishing, pp. 185–210. https://doi.org/10.4018/979-8-3693-9611-7.ch007

Alansary, S.A., Ayyad, S.M., Talaat, F.M., Saafan, M.M., 2025. Emerging AI threats in cybercrime: a review of zero-day attacks via machine, deep, and federated learning. Knowl Inf Syst. https://doi.org/10.1007/s10115-025-02556-6

Arpaci, I., 2024. A Multianalytical SEM-ANN Approach to Investigate the Social Sustainability of AI Chatbots Based on Cybersecurity and Protection Motivation Theory. IEEE Trans. Eng. Manage. 71, 1714–1725. https://doi.org/10.1109/TEM.2023.3339578

Azrour, M., Mabrouki, J., Guezzaz, A., Alabdulatif, A. (Eds.), 2025. AI-Enhanced Solutions for Sustainable Cybersecurity:, Advances in Information Security, Privacy, and Ethics. IGI Global. https://doi.org/10.4018/979-8-3693-6371-3

Belton, V., Stewart, T.J., 2002. Multiple Criteria Decision Analysis. Springer US, Boston, MA. https://doi.org/10.1007/978-1-4615-1495-4

Bhimavarapu, U., 2025. AI-Enhanced Green IT Security Bridging Cybersecurity and Sustainability:, in: Gupta, B.B., Pramod, D., Moslehpour, M. (Eds.), Advances in Computational Intelligence and Robotics. IGI Global, pp. 127–140. https://doi.org/10.4018/979-8-3693-8034-5.ch006

Brevini, B., 2020. Black boxes, not green: Mythologizing artificial intelligence and omitting the environment. Big Data & Society 7, 2053951720935141. https://doi.org/10.1177/2053951720935141

BustamanteV, D., Sánchez, L.E., Rosado, D.G., Santos-Olmo, A., Fernández-Medina, E., 2025. Towards a sustainable cybersecurity framework for Agriculture 4.0 based on a systematic analysis of proposals. Computers & Security 159, 104650. https://doi.org/10.1016/j.cose.2025.104650

Dawes, J., 2008. Do Data Characteristics Change According to the Number of Scale Points Used? An Experiment Using 5-Point, 7-Point and 10-Point Scales. International Journal of Market Research 50, 61–104. https://doi.org/10.1177/147078530805000106

De Vries, A., 2023. The growing energy footprint of artificial intelligence. Joule 7, 2191–2194. https://doi.org/10.1016/j.joule.2023.09.004

Dr. Varsha, P.S., 2023. How can we manage biases in artificial intelligence systems – A systematic literature review. International Journal of Information Management Data Insights 3, 100165. https://doi.org/10.1016/j.jjimei.2023.100165

Finstad, K., 2010. Response interpolation and scale sensitivity: Evidence against 5-point scales. Journal of usability studies.

Gómez, Á.L.P., Maimó, L.F., Celdrán, A.H., Clemente, F.J.G., 2023. SUSAN: A Deep Learning based anomaly detection framework for sustainable industry. Sustainable Computing: Informatics and Systems 37, 100842. https://doi.org/10.1016/j.suscom.2022.100842

Gregor, S., Hevner, A.R., 2013. Positioning and Presenting Design Science Research for Maximum Impact1. MIS Quarterly 37, 337–355. https://doi.org/10.25300/MISQ/2013/37.2.01

Gupta, B.B., Pramod, D., Moslehpour, M. (Eds.), 2025. Sustainable Information Security in the Age of AI and Green Computing:, Advances in Computational Intelligence and Robotics. IGI Global. https://doi.org/10.4018/979-8-3693-8034-5

Hamid, I., Rahman, M.M.H., 2025. AI, machine learning and deep learning in cyber risk management. Discov Sustain 6, 389. https://doi.org/10.1007/s43621-025-01012-3

Henao, F., Edgell, R., Sharma, A., Olney, J., 2025. AI in power systems: a systematic review of key matters of concern. Energy Inform 8, 76. https://doi.org/10.1186/s42162-025-00529-1

Hevner, A.R., March, S.T., Park, J., Ram, S., 2004. Design Science in Information Systems Research1. MIS Quarterly 28, 75–106. https://doi.org/10.2307/25148625

Keeney, R.L., Raiffa, H., 1993. Decisions with Multiple Objectives: Preferences and Value Trade-Offs, 1st ed. Cambridge University Press. https://doi.org/10.1017/CBO9781139174084

Königstorfer, F., Thalmann, S., 2022. AI Documentation: A path to accountability. Journal of Responsible Technology 11, 100043. https://doi.org/10.1016/j.jrt.2022.100043

Le Goff, T., 2025. Environmental law's principles applied to artificial intelligence: a path towards regulation. International Review of Law, Computers & Technology 1–24. https://doi.org/10.1080/13600869.2025.2506166

Lezzi, M., Montefusco, P., Lazoi, M., Corallo, A., 2025. AI-based cybersecurity for a sustainable digital industry: Systematic literature review and future research directions. Journal of Industrial Information Integration 48, 100980. https://doi.org/10.1016/j.jii.2025.100980

Mohamed, N., 2025. Artificial intelligence and machine learning in cybersecurity: a deep dive into state-of-the-art techniques and future paradigms. Knowl Inf Syst 67, 6969–7055. https://doi.org/10.1007/s10115-025-02429-y

Montibeller, G., Von Winterfeldt, D., 2015. Cognitive and Motivational Biases in Decision and Risk Analysis. Risk Analysis 35, 1230–1251. https://doi.org/10.1111/risa.12360

OECD, European Union, Joint Research Centre - European Commission, 2008. Handbook on Constructing Composite Indicators: Methodology and User Guide. OECD. https://doi.org/10.1787/9789264043466-en

Pendleton, M., Garcia-Lebron, R., Cho, J.-H., Xu, S., 2017. A Survey on Systems Security Metrics. ACM Comput. Surv. 49, 1–35. https://doi.org/10.1145/3005714

Priebe, T., Kieseberg, P., Adrowitzer, A., Eigner, O., Kovac, F., 2023. Sustainability Effects of Robust and Resilient Artificial Intelligence, in: Holzinger, A., Kieseberg, P., Cabitza, F., Campagner, A., Tjoa, A.M., Weippl, E. (Eds.), Machine Learning and Knowledge Extraction, Lecture Notes in Computer Science. Springer Nature Switzerland, Cham, pp. 188–199. https://doi.org/10.1007/978-3-031-40837-3_12

Rafique, S.H., Abdallah, A., Musa, N.S., Murugan, T., 2024. Machine Learning and Deep Learning Techniques for Internet of Things Network Anomaly Detection—Current Research Trends. Sensors 24, 1968. https://doi.org/10.3390/s24061968

Ranpara, R., Alsalman, O., Kumar, O.P., Patel, S.K., 2025. A simulation-driven computational framework for adaptive energy-efficient optimization in machine learning-based intrusion detection systems. Sci Rep 15, 13376. https://doi.org/10.1038/s41598-025-93254-4

Rawat, R., 2025. Optimizing ML models for cybercrime detection: balancing performance, energy consumption, and carbon footprint through multi-objective optimization. Discov Artif Intell 5, 38. https://doi.org/10.1007/s44163-025-00262-4

Shan, L., 2025. (IoT) Network intrusion detection system using optimization algorithms. Sci Rep 15, 21706. https://doi.org/10.1038/s41598-025-04638-5

Uddin, M., Irshad, M.S., Kandhro, I.A., Alanazi, F., Ahmed, F., Maaz, M., Hussain, S., Ullah, S.S., 2025. Generative AI revolution in cybersecurity: a comprehensive review of threat intelligence and operations. Artif Intell Rev 58, 236. https://doi.org/10.1007/s10462-025-11219-5

Usman, Y., Ihejirika, C.J., Offor, S.N., Akl, R., Chataut, R., 2025. Green Cybersecurity: Leveraging AI, ML, and LLMs to Optimize Energy, Threat Detection, and Sustainability Frameworks. IEEE Access 13, 159345–159379. https://doi.org/10.1109/ACCESS.2025.3602451

Yang, K., Wang, J., Li, M., 2024. An improved intrusion detection method for IIoT using attention mechanisms, BiGRU, and Inception-CNN. Sci Rep 14, 19339. https://doi.org/10.1038/s41598-024-70094-2

Yu, X., Fan, L., Yu, Y., 2025. Artificial Intelligence and Corporate ESG Performance: A Mechanism Analysis Based on Corporate Efficiency and External Environment. Sustainability 17, 3819. https://doi.org/10.3390/su17093819

Territorial Development as Resource for Firm Growth Strategies: A Conceptual Framework

Stefano Marciano, Raffaele Fiorentino, Rosita Capurro, Francesca Peluso
University of Naples Parthenope, Naples, Italy

Abstract

This study aims to investigate the role that territorial development can assume within firm growth strategies. Specifically, the study aims to provide an in-depth analysis of territorial development for firms' value creation processes and the achievement of the competitive advantage. In order to achieve this aim, the study analyze the literature on the Resource-Based View Theory, dynamic capabilities and territorial development. Thus, the study has developed a conceptual framework, grounded in the Resource-Based View Theory, aimed at clarifying the conditions under which territorial development can acquire strategic relevance within firm growth strategies. The findings support the idea that territorial resources become strategically relevant when they are aligned with the firm's objectives, business model and strategic trajectory; when they present the features of durability and limited transferability and replicability; and when the firm is able to capture the value generated. In this perspective, territorial development does not contribute to firm growth simply because some resources are available within a given territory. Its strategic value depends on the firm's capability to identify and activate such resources and transform them into a coherent basis for competitive advantage.

The paper provides relevant contributions in theory and practice. From a theoretical perspective, the results contribute to the strategic management literature by extending the Resource-Based View Theory to the interpretation of territorial development as a strategic resource for firm growth. From a practical perspective, the study offers a conceptual framework useful for entrepreneurs, executives, managers and policy-makers to identify, evaluate and exploit territorial resources. However, as a theoretical and literature-based study, the proposed framework requires empirical validation and further investigation across different sectors, territorial settings and firm types in order to assess its explanatory power and practical applicability in heterogeneous business contexts. Future studies should integrate the theoretical analysis with empirical investigations in specific territorial contexts.

Keywords: Territorial development; strategic management; Resource-Based View; firm growth.

Paper type: Academic Research Paper

1 Introduction

Territorial development has become an increasingly relevant issue in contemporary strategic and managerial debates, as it lies at the intersection of economic growth, sustainability, territorial planning and firms' competitiveness (Porter, 2000; Delgadillo et al., 2021; Morretta et al., 2021). From this perspective, territorial development can be understood as a multidimensional configuration of tangible and intangible resources, social relations, traditions and knowledge dynamics that shape value creation processes (Allais et al., 2015; Torre, 2025).

In many cases, firms are embedded in territorial contexts where access to human capital, infrastructures, local institutions, social capital and place-based knowledge affects both strategic choices and competitive trajectories. The access to these territorial resources often can generate significant competitive benefits, as it can enhance firms' value creation processes, support the achievement of competitive advantage and contribute to the development of more sustainable growth paths (Oerlemans and Meeus, 2005).

Although the growing interest devoted to this topic in the literature (Bianchi and Richiedei, 2023; Fiorentino, 2023; Fiume et al., 2025; Capurro et al., 2025), there has been limited attention given to interpreting territorial development through a strategic management lens, particularly as a resource for firm growth. Despite their potential strategic relevance, territorial resources are often under-exploited by firms, even when they can support the achievement of competitive advantage (Storper, 1997; Porter, 2000; Allais et al., 2015).

In order to bridge the gap, the study aims to investigate the role that territorial development can assume within firm growth strategies. More specifically, the study addresses the following research question: which is the role of territorial development for firm growth strategies from a Resource-Based View perspective? To answer this question, the study aims to provide an in-depth analysis of territorial development and advances a conceptual framework grounded in the Resource-Based View (RBV), with particular reference to Grant's resource and capability-based approach to strategy formulation (Grant, 1991). This perspective is especially suitable because it makes possible to move beyond a static interpretation of territorial resources and to focus

instead on the conditions under which they become strategically relevant, contribute to competitive advantage and enable the appropriation of value.

The paper is structured as follows. Section 2 presents the theoretical background. Section 3 develops the proposed conceptual framework. Section 4 discusses the main findings and implications emerging from the framework. Finally, Section 5 concludes the paper by summarizing the main contributions, outlining the practical implications and suggesting avenues for future research.

2 Theoretical Background

2.1 Resource-based view

The RBV has emerged as one of the most influential frameworks in strategic management for understanding how firms achieve and sustain competitive advantage. Originally developed by Barney (1991, 2001) and expanded by other scholars (Grant, 1991; Peteraf, 1993; Sirmon et al., 2007), RBV advances that a firm's competitive advantage derives from its resources and capabilities and from the way these are organized to support strategy formulation. While Barney's contribution has been pivotal in emphasizing the role of valuable, rare, inimitable and non-substitutable resources, Grant's perspective is particularly relevant in highlighting that strategy is grounded in the match between the firm's resources and capabilities and the opportunities arising in its external environment.

The evolution of RBV theory has been marked by increasing complexity in understanding resource heterogeneity and its relationship to performance, moving from static conceptions of resource possession to dynamic understandings of resource orchestration and reconfiguration (D'oria et al. 2021). The foundational premise of RBV emphasizes that competitive advantages arise primarily from internal resources and capabilities. This theoretical shift has profound implications for strategic management practice, redirecting managerial attention from competitive analysis to internal resource and capability development.

The integration of RBV with other theoretical perspectives, particularly Dynamic Capabilities Theory (Grant, 1991; Teece et al., 1997; Teece, 2007), has enriched our understanding of how firms must continuously adapt and reconfigure their resources to maintain competitive advantages in dynamic environments.

Dynamic Capabilities Theory addresses a critical limitation of traditional RBV by explaining how firms maintain competitive advantage in rapidly changing environments. While RBV emphasizes the possession of valuable resources, dynamic capabilities focus on the firm's ability to identify opportunities and threats, exploit promising opportunities and transform resource configurations in response to environmental changes.

Firms that successfully combine possession of resources with robust dynamic capabilities achieve superior and more sustainable performance outcomes (Wilden et al., 2016). This perspective has particular relevance for understanding how firms can leverage territorial development.

In line with the RBV and dynamic capabilities perspective, territorial development cannot be interpreted as a static resource; rather, it represents an evolving configuration that requires active management, continuous strategic scanning of emerging territorial opportunities and the purposeful reconfiguration of firm–territory engagement strategies.

2.2 Territorial development

The concept of territorial development exhibits a broad and multidimensional nature that encompasses economic, social, environmental and institutional components (Allais et al., 2015). As it has been defined by Torre (2025), territorial development can be understood as a process that aims to improve the well-being and wealth of a territory's stakeholders through cooperation and competition, collective initiatives, governance processes and territorial innovation dynamics. In this sense, territory is a relational space in which a community recognises itself and uses it as a reference point for its individual and collective actions, contributing to the orientation of its development trajectories (Maillat, 1995; Banini et al., 2021). Consistently, territory can also be understood as an evolving and complex combination of actors and geographical space, shaped by the ways in which such actors use, organise and manage it (Allais et al., 2015; Moine, 2006).

In this sense, development is not solely grounded in sectoral dynamics, but also in the territory as a context for coordination, learning and collective action, where the valorisation of local resources and the construction of social capital play a decisive role (Shucksmith, 2010). More precisely, social capital, defined as a set of trust, reciprocity and networks of cooperation, is a vital territorial resource because it enhances the territory's ability to effectively utilise its material and immaterial resources (Portes, 2000; Tura and Harmaakorpi, 2005). In addition, territorial value can be seen as arising from the interaction among multiple interconnected

components, including natural ecosystems, anthropic infrastructures, industrial systems and social spaces, each of which provides both tangible and intangible resources (Allais et al., 2015; Moine, 2006). In particular, social spaces contribute with elements such as common culture, local history and identity, all of which can shape how territorial resources are perceived, mobilised and transformed into value (Allais et al., 2015; Moine, 2006).

This leads to a strategic interpretation of territory as a driver of economic and entrepreneurial development, in which firms, resources and knowledge are closely intertwined and embedded with the territorial context (Darroch, 2005). From this view, territorial characteristics can affect firm performance. This suggests that territorial capital can be interpreted as a resource for competitive advantage of firms (Morretta et al., 2020).

Moreover, territorial resources are often under-exploited by firms, despite their potential to generate competitive advantage through the use of local tangible and intangible resources, such as specialised competences and excellence clusters (Porter, 2000; Allais et al., 2015).

2.3 Territorial development from a RBV perspective

The conceptualization of territorial development as a resource reflects an important theoretical extension of RBV that recognizes the strategic importance of local and regional contexts. Geographic proximity creates both tangible advantages - such as reduced transportation and logistics costs, improved supply chain coordination, access to specialized labour markets, shared infrastructure and proximity to suppliers and key inputs - and intangible benefits including knowledge, institutional support and cultural alignment (Oerlemans and Meeus, 2005). Research on industrial clusters demonstrates that geographic concentration of firms within specific industries creates unique competitive environments where territorial factors significantly enhance firm performance and innovation capacity (Storper, 1997; Porter, 2000).

Contrasting passively accepting a geographic location, strategic territorial development involves active engagement with regional economies, community stakeholders and institutional actors to enhance the strategic value of territorial resources (Morretta et al., 2020). This management approach recognizes that territorial advantages require continuous cultivation; firms have to invest in building relationships, improving infrastructure, developing knowledge networks and engaging with institutions to fully realize the value potential of territorial resources.

Territorial development management has evolved as a central strategic function for firms operating in sectors strongly tied to place-based resources, such as tourism and creative industries. In such contexts, firms are actors embedded within geographically bounded systems of relationships, where performance depends on the ability to access, combine and leverage territorial resources and capabilities (Boschma, 2005). Effective territorial management involves orchestrating relationships between multiple stakeholders' groups to create an environment that fosters success (Freeman, 1984).

3 Conceptual Framework

In light of the theoretical background, this study proposes a conceptual framework designed to interpret territorial development within the RBV, with particular reference to Grant's resource and capability-based approach to strategic management (Grant, 1991). The framework was developed through literature analysis and a process of synthesis aimed at identifying the key concepts and theoretical relationships emerging from RBV, dynamic capabilities and territorial development (Capurro et al., 2024).

Territorial resources do not automatically constitute strategic assets; their strategic relevance depends on the firm's ability to recognize their meaning, evaluate their potential and effectively integrate them with its internal resource and dynamic capabilities.

The proposed framework is grounded in a fundamental distinction between territorial resources and firm dynamic capabilities. The former encompasses the set of tangible and intangible resources embedded within a given territorial context, whereas the latter refer to the firm's ability to access, coordinate and integrate such resources with its internal resource and dynamic capabilities (Barney, 1991; Grant, 1991, Teece 1997; 2007).

The achievement of competitive advantage depends on the firm's ability to translate access to territorial resources into an effective basis for strategic action. Territorial development does not constitute a capability per se; it represents a pool of potentially strategic resources whose value is realized only when the firm develops the capabilities of absorption, integration, coordination and orchestration. The potential value-creation of territorial development can be examined along three closely interrelated dimensions: their capacity to generate competitive advantage, their ability to sustain such competitive advantage over time and their effectiveness in enabling the appropriation of the value created (Grant, 1991).

The first dimension advances that a resource gets strategic significance when it exhibits two fundamental characteristics: relevance to the market's critical success factors and scarcity. From a territorial development

perspective, a territorial resource becomes strategic insofar as it concretely contributes to the achievement of the firm's objectives. Accordingly, not all territorial resources can be considered, per se, as sources of competitive advantage, only those resources that are closely aligned with the firm's business model and strategic trajectory assume strategic relevance. This is the case, for instance, with a local pool of highly skilled labour, place-based know-how developed within a specific production tradition, close access to specialised suppliers and service providers, supportive local institutions, dense networks of trust and collaboration and a territorial reputation that enhances the firm's legitimacy and market positioning (Darroch, 2005; Allais et al., 2015; Moine, 2006; Torre, 2025).

Alongside strategic relevance, scarcity assumes a distinctive meaning. It should be interpreted primarily in terms of differential accessibility. Territorial resources are inherently location-specific and unequally distributed across space; however, their scarcity also derives from the fact that, although potentially accessible by multiple actors, access is not equally guaranteed to all firms operating within the same context. The ability to effectively leverage such resources depends on the firm's degree of local embeddedness, its legitimacy and its capacity to integrate into local networks of cooperation (Tura and Harmaakorpi, 2005).

The second dimension of the framework concerns the sustainability of competitive advantage over time. Such sustainability depends on three key properties: durability, transferability and replicability.

The durability of territorial resources comes from their origin in long-term historical, cultural and institutional processes. Social capital, territorial reputation, productive traditions, the quality of relationships between firms and institutions and the accumulation of knowledge are all elements that develop gradually and, for this reason, tend to exhibit strong persistence over time (Portes 2000, Allais et al., 2015; Moine, 2006). This implies that competitive advantages grounded in such resources can present greater resilience compared to those based on market-acquired assets.

The limited transferability of territorial resources further reinforces this advantage. Their strategic relevance lies, in part, in their location-specific nature, which makes them difficult to relocate or transfer without a significant loss of value. In the case of territorial resources, this constraint is even more pronounced as they depend on a contextual configuration that cannot be easily reproduced elsewhere. While firms can be able to relocate certain tangible assets, they cannot transfer with the same effectiveness the complex set of relationships and coordination mechanisms that underpin the strategic value of a given territorial context.

This is completed by the limited replicability of territorial resources. Even when competitors are able to observe the factors that make a territory advantageous, reproducing the same configuration in a different context remains highly challenging. Territorial resources do not derive from a single component; instead, they emerge from an intricate combination of material and immaterial, relational and institutional, formal and informal elements. Such complexity makes imitation difficult, as what ultimately matters is not the presence of individual assets, but the way in which they are combined and mutually reinforced over time. Consequently, competitive advantage lies in the firm's ability to operate within a complex territorial configuration whose replication by competitors faces substantial barriers.

The third dimension concerns the appropriation of the value generated. The availability of a resources does not automatically imply the ability to capture the returns that it produces. This issue becomes even more critical in the case of territorial resources, which are typically not directly owned by the firm. Elements such as the skills diffused within a community, reputation, relationships with local institutions, availability of shared knowledge and social capital of a territory constitute resources to which firms can gain access, but which they rarely full control.

Therefore, the strategic challenge lies in accessing these resources and transforming such access into economic, competitive and reputational advantages and, more broadly, into value (Gereffi et al., 2005; Darroch, 2005). The appropriation of territorial value depends, first, on the presence of appropriability mechanisms, specifically in formal and informal arrangements capable of strengthening the link between the value generated within the territory and the firm's performance. These mechanisms can take the form of contractual agreements, partnerships, specific investments or territorial branding initiatives.

Second, the firm's relative bargaining power within the territorial system plays a crucial role (Porter, 1979). Firms operating within the same territory do not possess equal capacity to take the benefits arising from territorial resource. This capacity tends to increase when a firm occupies central positions within local networks, controls critical nodes of the value chain, possesses distinctive capabilities complementary to territorial resources or enjoys a high degree of legitimacy among key stakeholders. Finally, the degree of territorial embeddedness is a key determinant.

The more deeply a firm is embedded in local economic, social and institutional networks, the greater its access to resources and its capacity to transform that access into relationships, reputation and economically

returns. From this perspective, embeddedness operates both upstream, as a condition enabling access to territorial resources and downstream, as a factor influencing the firm's capacity to capture the value generated.

The conceptual framework is presented in Figure 1.

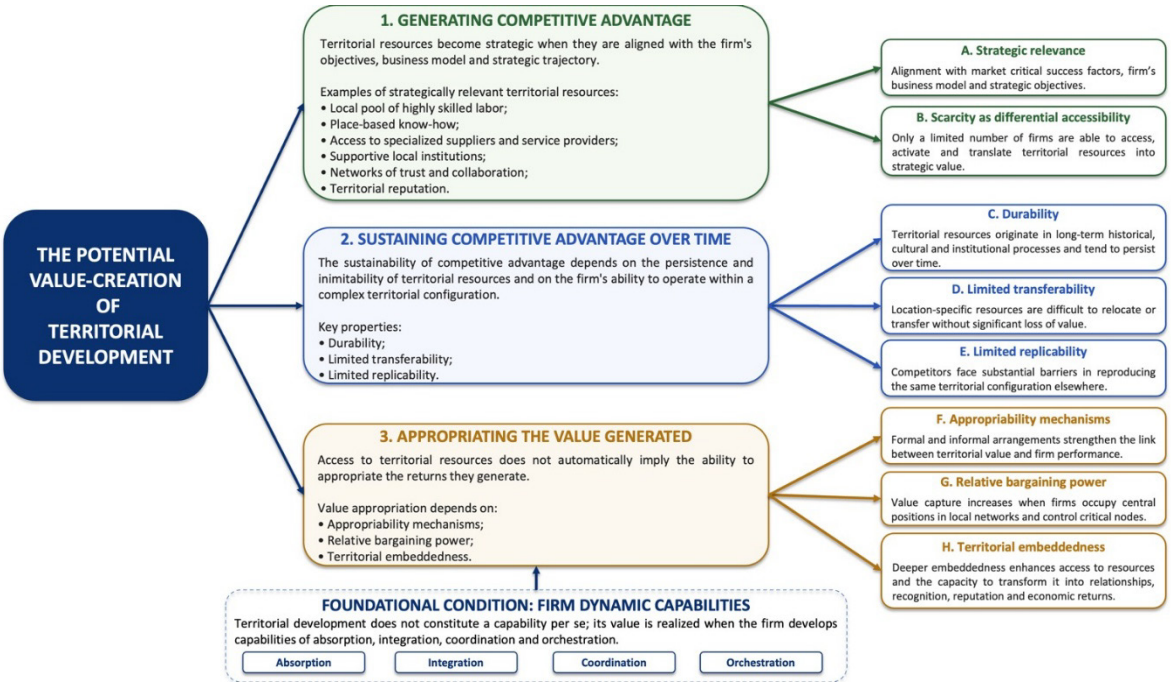


Figure 1: The potential value-creation of territorial development: a resource-based framework

4 Discussion

From the elaboration of the conceptual framework, several findings emerge regarding the role of territorial development in firm's growth strategy.

Primarily, the analysis shows that territorial development cannot be interpreted as an external condition that is generally favourable to the firm, nor as a capability per se; it should be seen as a set of potentially strategic territorial resources, whose value is neither intrinsic nor automatic, but depends on the firm's ability to recognize their relevance, assess their potential and integrate them with its own resource and dynamic capabilities. In this regard, the primary finding of this study is to shift the focus from the territory as a contextual backdrop to the territory as a source of resources that can be strategically exploited.

With regard to the nature of the relationship between the firm and its territorial context, the framework highlights that competitive advantage does not arise from the availability of local resources - since these are, in principle, accessible to all firms operating within the same area - but from the firm's ability, first, to effectively access such resources and, subsequently, to transform them into strategic resources.

This step is particularly relevant, as it enables a clearer distinction between local endowments and strategic value. The presence within a given context of human capital, knowledge, infrastructure, supporting institutions and networks of cooperation does not automatically translate into competitive advantage. The framework emphasizes that such resources become strategically relevant only when they are aligned with the firm's objectives, its business model and its strategic trajectory. Therefore, the strategic value of a territory is often selective: it depends on what the territory "possesses," and on the degree of alignment between the territory's characteristics and the firm's competitive capabilities. The concept of scarcity is closely linked with this perspective. The framework demonstrates that, in this context, scarcity should be understood primarily as differential accessibility. Territorial resources are inherently location-specific and unevenly distributed; however, their strategic relevance also derives from the fact that, even when they are embedded within the same geographical context, not all firms are equally able to access, activate and transform them into value. Territorial scarcity, therefore, depends on the firm's degree of embeddedness, its legitimacy and its capacity to integrate into local networks of cooperation. This represents a theoretically significant insight, as it enables a shift beyond a strictly proprietary interpretation of the RBV, highlighting that, in the case of territorial resources, competitive advantage is grounded in mechanisms of access and activation. In this perspective, scarcity coincides with the fact that only a limited number of firms are able to effectively exploit the territorial resources.

With regard to the sustainability of competitive advantage over time, the requirements of durability, transferability and replicability highlight that territorial resources can serve as a particularly robust foundation for sustainable advantage. Their durability derives from their embedding in long-term historical, cultural and institutional processes, while their limited transferability and replicability stem from their inherently territorial nature, which renders them difficult to relocate and challenging to reproduce. Therefore, the competitive advantage grounded in territorial resources can exhibit greater resilience than advantages based on assets that can be acquired on the market. From this perspective, the most significant insight is that the sustainability of such advantage resides in the overall territorial configuration and in the firm's ability to effectively operate within it.

Another key finding concerns value capture. The framework shows that, in the case of territorial resources, the strategic challenge extends beyond access to resources to the firm's ability to capture the returns generated by their use. This is a crucial point, since many territorial resources are not directly owned by the firm. Value capture depends on the presence of appropriability mechanisms, the firm's relative bargaining power within the territorial system and its degree of embeddedness in local networks. Consequently, embeddedness operates as a key factor shaping the firm's ability to transform such access into stable relationships, recognition, reputation and economic returns. From this perspective, the territory tends to reward firms that occupy central positions within networks, control critical nodes of the value chain, possess complementary capabilities and enjoy higher levels of legitimacy among key stakeholders.

These findings lead to understand that the contribution of territorial development to firm growth derive from the firm's ability to develop a set of dynamic capabilities oriented toward the absorption, integration, coordination and orchestration of territorial resources.

5 Conclusions

This study investigates the role that territorial development can assume for firms' growth strategies. Starting from the limited attention in the literature to territorial development from a RBV perspective, the paper developed a conceptual framework to clarify under which conditions territorial resources can acquire strategic relevance and contribute to the generation of sustainable competitive advantage and to the firm growth.

The findings make it possible to answer the research question by showing that territorial development acquires strategic relevance only under specific conditions. More precisely, territorial resources become a potential source of firm growth when they are consistent with the firm's objectives, business model and strategic trajectory and when the firm is able to appropriate the value generated. In this perspective, the framework clarifies that the contribution of territorial development to firm growth derive from firm's ability to recognize their strategic relevance, to access and activate them selectively and to integrate them with its internal resources and dynamic capabilities. From this perspective, territorial development emerges as a source of value and competitive advantage, whose impact on growth derives primarily from the firm's capacity for absorption, integration, coordination and orchestration.

The paper provides contributions to theory and practice.

From a theoretical perspective, the results contribute to the strategic management literature streams, by showing that territorial resources, although external to the firm in a strict ownership sense, can be incorporated into the firm growth strategy when they are strategically activated and combined with internal capabilities. In addition, it contributes to the debate on the relationship between firms and territory by clarifying that competitive advantage is not primarily grounded in the ownership of territorial resources, but in the firm's ability to access and exploit them selectively through embeddedness, legitimacy and network positioning.

Although the paper is mainly theoretical, it also offers some practical implications. The advanced framework can support entrepreneurs, managers and policy-makers in interpreting territorial development as a potential strategic driver. More specifically, the framework suggests that firms should identify which territorial resources are most consistent with their strategic trajectory, assess the conditions under which such resources can generate competitive advantage and strengthen the capabilities required to absorb, integrate, coordinate and orchestrate them. Moreover, it provides an initial analytical tool for interpreting territorial development within strategic decision-making processes.

Thus, the study presents some limitations that should be acknowledged. The paper does not include an empirical validation of the proposed framework. As a consequence, the identified relationships between territorial resources, dynamic capabilities and firm growth remain at a theoretical level and require further testing in real firm contexts.

The paper opens future research opportunities. The advanced framework should be empirically tested through further investigations aimed at verifying how firms actually recognize and exploit territorial resources

within their growth strategies. In addition, future studies would apply the framework to concrete territorial contexts in order to explore how place-based resources, institutional settings and local networks interact in shaping competitive advantage. In this regard, within a regional based research project founded by "Regione Campania" two Italian promising empirical settings concern the "Agro Nocerino-Sarnese" and the "Valle dell'Ufita", both of which represent local value chains where territorial development can be observed as a strategic resource in relation to distinctive productive, institutional, and socio-economic dynamics.

6 Ethics Declaration

Ethical clearance was not required.

7 AI Declaration

Artificial intelligence tools were not used in the preparation of this paper.

References

Allais, R., Reyes, T. and Roucoules, L. (2015) "Inclusion of Territorial Resources in the Product Development Process", *Journal of Cleaner Production*, Vol. 94, pp. 187–197.

Banini, T., Ilovan, O.R. and Paasi, A. (2021) *Representing Place and Territorial Identities in Europe*, Springer International Publishing.

Barney, J. (1991) "Firm Resources and Sustained Competitive Advantage", *Journal of Management*, Vol. 17, No. 1, pp. 99–120.

Barney, J. (2001) "Resource-Based Theories of Competitive Advantage: A Ten-Year Retrospective on the Resource-Based View", *Journal of Management*, Vol. 27, No. 6, pp. 643–650.

Bianchi, S. and Richiedei, A. (2023) "Territorial Governance for Sustainable Development: A Multi-Level Governance Analysis in the Italian Context", *Sustainability*, Vol. 15, No. 3, Article 2526.

Boschma, R. (2005) "Proximity and Innovation: A Critical Assessment", *Regional Studies*, Vol. 39, No. 1, pp. 61–74.

Capurro, R., Fiorentino, R., Garzella, S. and Marciano, S. (2024) "Metaverse: Value Creation and Performance Measurement Across Physical and Digital World", *Management Control*, No. 1.

Capurro, R., Garzella, S., Fiorentino, R. and Peluso, F. (2025) "Settore del Corallo e Sviluppo Territoriale: Evoluzione delle Reti Economiche e Creazione di Valore Allargato", *Rivista Italiana di Ragioneria e di Economia Aziendale*, No. 7–8, pp. 165–189.

D'Oria, L., Crook, T.R., Ketchen Jr, D.J., Sirmon, D.G. and Wright, M. (2021) "The Evolution of Resource-Based Inquiry: A Review and Meta-Analytic Integration of the Strategic Resources–Actions–Performance Pathway", *Journal of Management*, Vol. 47, No. 6.

Darroch, J. (2005) "Knowledge Management, Innovation and Firm Performance", *Journal of Knowledge Management*, Vol. 9, No. 3, pp. 101–115.

Delgadillo, E., Reyes, T. and Baumgartner, R.J. (2021) "Towards Territorial Product-Service Systems: A Framework Linking Resources, Networks and Value Creation", *Sustainable Production and Consumption*, Vol. 28, pp. 1297–1313.

Fiorentino, R., Landriani, L., Lardo, A. and Marciano, S. (2023) "Governance Systems for Civic Wealth Creation through a New Accessibility to Cultural Heritage: The Case of 'La Paranza'", in Aureli, S., Del Baldo, M., Demartini, P. and Piber, M. (eds.), *Cultural Heritage as a Trigger for Civic Wealth Creation and Sustainable Urban Development*, RomaTre-Press, pp. 33–56.

Fiume, R., Landriani, L., Lardo, A. and Marciano, S. (2025) "Creazione di Valore e Governance del Patrimonio Culturale: Il Ruolo del Capitale Umano nella Sostenibilità", *Rivista Italiana di Ragioneria e di Economia Aziendale*, No. I, pp. 6–38.

Freeman, R.E. (1984) *Strategic Management: A Stakeholder Approach*, Pitman, Boston, MA.

Grant, R.M. (1991) "The Resource-Based Theory of Competitive Advantage: Implications for Strategy Formulation", *California Management Review*, Vol. 33, No. 3, pp. 114–135.

Maillat, D. (1995) "Territorial Dynamic, Innovative Milieus and Regional Policy", *Entrepreneurship & Regional Development*, Vol. 7, No. 2, pp. 157–165.

Moine, A. (2006) "The Territory as a Complex System: An Operational Concept for Land Planning and Geography", *Espaces Géographiques*, Vol. 35, No. 2, pp. 115–132.

Morretta, V., Syrett, S. and Ramirez, L.S. (2020) "Territorial Capital as a Source of Firm Competitive Advantage: Evidence from the North and South of Italy", *European Planning Studies*, Vol. 28, No. 12, pp. 2390–2408.

Oerlemans, L. and Meeus, M. (2005) "Do Organizational and Spatial Proximity Impact on Firm Performance?", *Regional Studies*, Vol. 39, No. 1, pp. 89–104.

Peteraf, M.A. (1993) "The Cornerstones of Competitive Advantage: A Resource-Based View", *Strategic Management Journal*, Vol. 14, No. 3, pp. 179–191.

Porter, M.E. (1979) "How Competitive Forces Shape Strategy", *Harvard Business Review*, Vol. 57, No. 2, pp. 137–145.

Porter, M.E. (2000) "Locations, Clusters and Company Strategy", in Clark, G., Feldman, M. and Gertler, M. (eds.), *The Oxford Handbook of Economic Geography*, Oxford University Press, pp. 253–274.

Portes, A. (2000) "The Two Meanings of Social Capital", *Sociological Forum*, Vol. 15, No. 1, March, pp. 1–12.

Shucksmith, M. (2010) “Disintegrated Rural Development? Neo-Endogenous Rural Development, Planning and Place-Shaping in Diffused Power Contexts”, *Sociologia Ruralis*, Vol. 50, No. 1, pp. 1–14.

Sirmon, D.G., Hitt, M.A. and Ireland, R.D. (2007) “Managing Firm Resources in Dynamic Environments to Create Value: Looking Inside the Black Box”, *Academy of Management Review*, Vol. 32, No. 1, pp. 273–292.

Storper, M. (1997) *The Regional World: Territorial Development in a Global Economy*, Guilford Press.

Teece, D.J., Pisano, G. and Shuen, A. (1997) “Dynamic Capabilities and Strategic Management”, *Strategic Management Journal*, Vol. 18, No. 7, pp. 509–533.

Teece, D.J. (2007) “Explicating Dynamic Capabilities: The Nature and Microfoundations of (Sustainable) Enterprise Performance”, *Strategic Management Journal*, Vol. 28, No. 13, pp. 1319–1350.

Torre, A. (2025) “Contribution to the Theory of Territorial Development: A Territorial Innovations Approach”, *Regional Studies*, Vol. 59, No. 1, pp. 2193–2218.

Tura, T. and Harmaakorpi, V. (2005) “Social Capital in Building Regional Innovative Capability”, *Regional Studies*, Vol. 39, No. 8, pp. 1111–1125.

Wilden, R., Devinney, T.M. and Dowling, G.R. (2016) “The Architecture of Dynamic Capability Research Identifying the Building Blocks of a Configurational Approach”, *Academy of Management Annals*, Vol. 10, No. 1, pp. 997–1076.

Measuring the Unmeasurable? A Systematic Literature Review of University Technology Transfer Efficiency

Miklós Verseghi-Nagy, András Nemeslaki
Budapest University of Technology and Economics, Budapest, Hungary

Abstract

Innovation has become central to achieving high-value growth, particularly in economies seeking to escape the middle-income trap where competitiveness still rests on low wages and process cost-cutting. In triple helix innovation ecosystems, technology transfer is seen as a crucial link enabling public–private value co-creation. Yet the efficiency of these processes remains poorly understood. University technology transfer efficiency (UTTE) — the relationship between inputs devoted to transferring research into commercial applications and resulting outputs — poses significant conceptual and measurement challenges. Inputs are ambiguous and multidimensional (research expenditure, technology transfer office resources, organizational capabilities); outputs are heterogeneous (patents, licenses, spin-offs, licensing income, broader impacts); and no consensus exists on which efficiency concept should formally link them. Despite a substantial body of empirical literature, UTTE measurement remains fragmented across incompatible units, anchored to transactional proxies, and disconnected from the economic outcomes that justify the investment. What is being measured is not what matters; what matters is not being measured.

This paper investigates whether the existing literature offers an adequate conceptual and empirical basis for measuring UTTE — and if not, what is missing. Our systematic review examines 71 peer-reviewed studies published between 2006 and 2026, analyzing: (a) how inputs and outputs are defined and operationalized; (b) which methodological frameworks relate inputs to outputs; (c) how studies address data limitations and cross-institutional comparability; and (d) to what extent emerging metric frameworks from professional associations (e.g., ASTP) are reflected in recent empirical work.

We find that DEA-type non-parametric frontiers dominate methodologically, with stochastic frontier analysis providing a parametric alternative; that data limitations severely constrain cross-institutional and cross-country comparability; and that emerging professional frameworks have only partially penetrated empirical practice. Taken together, these findings reveal a field built around what is measurable and institutionally available rather than what is theoretically adequate — and point toward a reoriented research agenda that moves upward in aggregation level, from institutions to ecosystems; forward in the value chain, from transfer outputs to enterprise outcomes; and deeper in explanatory mechanism, from quantitative proxies to qualitative process understanding.

Keywords: Social Impact Measurement, University Technology Transfer Efficiency

Paper type: Academic Research Paper

1 Research methodology: protocol and reporting

The review includes peer-reviewed journal articles published between 2006 and 2026 that: conceptualize UTTE as an explicit input-output transformation (e.g., via DEA, frontier models, or maturity frameworks); use universities, groups of universities, or TTOs as the unit of analysis; explicitly identify both input variables (e.g., R&D expenditure, staff, TTO resources) and output variables (e.g., patents, licenses, spin-offs, licensing income); and apply a quantitative efficiency framework or clearly specified perceptual constructs of TTO performance. Studies were excluded if they described technology transfer outcomes without an explicit efficiency framework, or focused solely on determinants of performance without measuring the input-output relationship itself.

Web of Science Core Collection and Scopus were searched using two concept dimensions combined in titles, abstracts, and keywords: university technology transfer ("university technology transfer" OR "technology transfer office") and efficiency ("efficiency" OR "productivity" OR "performance"). Only English-language peer-reviewed articles were included. Initial searches yielded 253 unique records; after title/abstract screening and independent crosschecking against eligibility criteria, 71 articles were retained. Data extraction followed PRISMA 2020 requirements, covering UTTE definitions, input–output configurations, and methodological approaches.

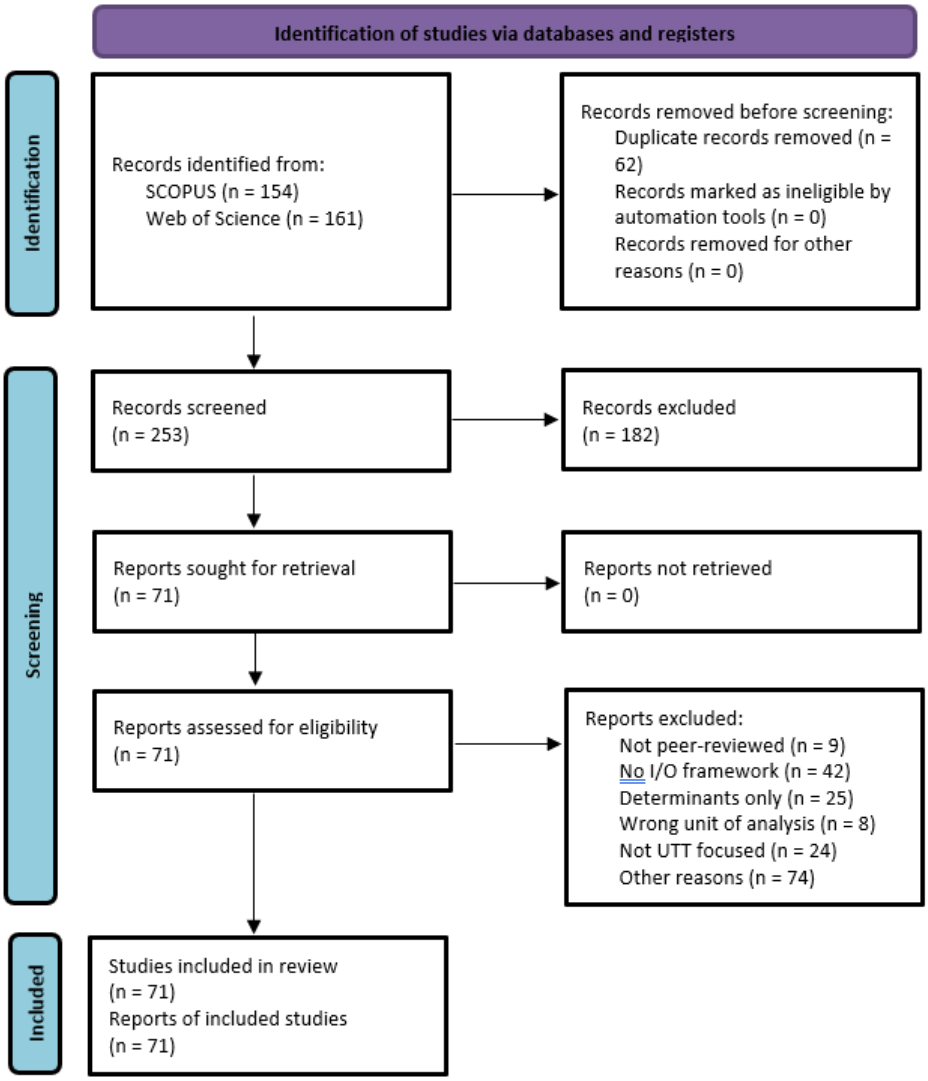


Figure 1: Identification of studies via databases and registers

2 Results

Methodologically, the field's foundations rest on early stochastic frontier and non-parametric analyses of TTO efficiency (Siegel, Waldman and Link, 2003; Chapple et al., 2005), the latter offering one of the first systematic parametric–non-parametric comparisons for UK universities. Within the review period, Anderson et al. (2007) refined DEA by introducing weight restrictions for benchmarking US institutions, Curi et al. (2012) applied DEA-based efficiency analysis to French universities, Bradley et al. (2013) synthesised university technology transfer models and empirical methods, and Lafuente and Berbegal-Mirabent (2019) incorporated portfolio complexity into TTO productivity analysis.

Across the 71 reviewed studies, four structural patterns emerge — in how inputs and outputs are modelled, in the methodological frameworks applied, in how data limitations are managed, and in the geographical scope of the evidence base.

3 Discussion

The UTTE literature has not yet produced a coherent basis for measuring university technology transfer efficiency. Four structural dimensions point toward a common problem: UTTE research has been built around what is measurable and institutionally available rather than what is theoretically adequate. Patents, licenses, and spin-off counts are tractable; the economic value they generate is not. DEA frontiers are rigorous; the efficiency scores they produce are local and context-specific. The sections that follow discuss each of these features in turn — not only to establish what is missing, but why it is missing. The common thread is a

transactional paradigm: the literature models technology transfer as a unilateral production process rather than the reciprocal, translational interaction that Triple Helix theory describes.

3.1 How inputs and outputs are currently modeled

The reviewed studies converge on three input and four output dimensions.

Common input dimensions

- Research capacity: R&D expenditure, R&D personnel, research intensity (Hülsbeck et al., 2013; Xia et al., 2020; Barra and Zotti, 2018).
- TTO resources/capabilities: staff numbers, experience, seniority, organizational structure, embeddedness, IP management capabilities (Caldera and Debande, 2010; Cartaxo and Godinho, 2017; Prónay et al., 2022; Secundo et al., 2017; Hülsbeck et al., 2013; Conti and Gaulé, 2011; Feng et al., 2012; Heisey and Adelman, 2011; Fernández-Alles et al., 2025; Iacobucci et al., 2021; Modic and Suklan, 2022; Sallan and Lordan, 2025).
- Intellectual capital: human, structural and relational capital of universities and TTOs (Secundo et al., 2017; Feng et al., 2012; Kreiling and Bounfour, 2020).

Common output dimensions

- IP & deals: invention disclosures, patent applications/grants, licenses, options, license income (Caldera and Debande, 2010; Lafuente and Berbegal-Mirabent, 2019; Hülsbeck et al., 2013; Vinig and Lips, 2015; Curi et al., 2012; Heisey and Adelman, 2011; Anderson et al., 2007; Tseng and Raudensky, 2014a; Tseng and Raudensky, 2014b; Showalter and Jensen, 2019).
- New firms: spin-offs/startups with/without university equity (Fadeyi et al., 2019; Lafuente and Berbegal-Mirabent, 2019; Kim, 2013; Vinig and Lips, 2015; Djokovic and Souitaris, 2008; Iacobucci et al., 2021; Lee and Jung, 2024; Lynskey, 2008).
- Broader commercialization: contract research, collaborative projects, consulting, science park benefits, job creation (Caldera and Debande, 2010; Fadeyi et al., 2019; Kim, 2013; Wang and Zhao, 2022; Kurman, 2011; Weis et al., 2018; Villasalero, 2014; Bigliardi et al., 2015; Li and Tan, 2020; Pujotomo et al., 2025).
- Perceptual/intangible performance: academics' willingness to cooperate, industry engagement, perceived TTO effectiveness (Prónay et al., 2022; Secundo et al., 2017; Secundo et al., 2019; Stankevičienė et al., 2017).

Three structural problems follow from this configuration. First, relying on patents and licenses as primary outputs risks introducing a disciplinary bias: science and engineering universities may appear more efficient partly because their outputs are more patent-intensive, not necessarily because they transfer more knowledge overall (Caldera and Debande, 2010; Curi et al., 2012). Second, organizational capabilities — embeddedness, networking depth, TTO maturity — shape observed efficiency yet are rarely measured directly, meaning that institutions pursuing non-IP channels are penalized by metrics that cannot capture those choices (Cartaxo and Godinho, 2017; Prónay et al., 2022; Secundo et al., 2017). Third, highly skewed distributions mean that DEA best-practice frontiers are often defined by a few structurally exceptional institutions, against which the majority are benchmarked (Kim, 2013; Kurman, 2011; Weis et al., 2018). Together, these problems suggest the input-output framework generates precise but potentially misleading efficiency assessments for much of the university population.

3.2 Main methodological frameworks linking inputs, outputs and efficiency

Across UTTE studies, non-parametric frontier methods dominate, complemented by parametric frontiers, index and multicriteria approaches. The choice of framework shapes what "efficiency" means in practice: frontier models yield relative rankings against a best-practice envelope, while index and structural approaches offer more diagnostic insight into specific performance drivers.

Frontier efficiency / productivity models

DEA dominates, modelling TTOs and universities as multi-input–multi-output production units from which relative efficiency scores are derived (Anderson et al., 2007; Kim, 2013; Wang and Zhao, 2022; Curi et al., 2012; Curi et al., 2015; Ma et al., 2022; Huang et al., 2011; Chen, K. et al., 2024a; Chen, X. et al., 2024b; See et al., 2023; Liu et al., 2026; Zarea et al., 2025; Lee et al., 2023). Variants — CCR/BCC, Slack-Based Measure, and Malmquist indices — extend the model to handle returns to scale, slack analysis, and productivity change over time. SFA offers a parametric alternative separating inefficiency from statistical noise, though far less used (Juárez and Montoro-Sánchez, 2019; Coupet and Ba, 2022; Siegel et al., 2008; Lee and Jung, 2021; Barra and Zotti, 2018).

The dominance of DEA is consequential: it yields relative rankings within a sample, not an absolute benchmark — leaving cross-system comparison structurally constrained.

Index and meta-data approaches

Index-based "health" measures: h-index–like commercialization, job-creation, and licensing-speed indices capturing skewed outcome distributions and time dynamics rather than simple counts (Kurman, 2011). Meta-data potential vs. realized transfer: uses research output to estimate a university's potential for transfer, then relates this to actual patents, licenses and spin-offs to derive a performance ratio (Vinig and Lips, 2015).

Multicriteria decision & structural models

Multicriteria decision methods (FARE + TOPSIS): build weighted composite performance scores from multiple TT indicators (e.g., patents, contracts, spin-offs) and rank universities (Stankevičienė et al., 2017; Belgin et al., 2025; Ustundag et al., 2011). Structural Equation Modeling / PLS-SEM: relate TTO resources and activities to multiple performance outputs via latent constructs rather than an efficiency frontier (Cartaxo and Godinho, 2017; Hwang et al., 2023).

The following table summarizes the main quantitative approaches, and the studies in which they occur.

Table 1: Main quantitative approaches to UTTE assessment

Approach	What it measures	Example use	Citations
DEA / SBM / Malmquist / SFA	Non-parametric and parametric efficiency frontiers	Nigeria, France, US, UK, China, TTO benchmarking	Fadeyi et al., 2019; Lafuente and Berbegal-Mirabent, 2019; Kim, 2013; Sutopo et al., 2019; Sutopo et al., 2022; Wang and Zhao, 2022; Curi et al., 2012; Siegel et al., 2008
Index-based (h-index analogues)	"Health" of commercialization, job creation, speed	US institutions	Kurman, 2011; Weis et al., 2018
Meta-data potential vs. realized transfer	Normalizes outputs by research-output-based potential	Dutch universities	Vinig and Lips, 2015
Maturity / capability models	TTO process maturity and capability as efficiency drivers	European TTOs	Kreiling and Bounfour, 2020; Scanlan, 2018
Configurational / fsQCA	Causal configurations for high transfer performance	China	Wang, 2025; Huang et al., 2025; Yang and Chen, 2025

3.3 Data limitations and cross-institution comparability

Across UTTE research, data are often incomplete, non-standardized, and context-specific, and most studies only partially address these constraints. The tension between internal validity (restricting to comparable units) and external validity (generalizing across systems) runs through virtually all empirical UTTE work.

Three coping strategies recur: harmonizing indicators via shared survey templates (AUTM or national equivalents) to stabilize definitions within a country (Kim, 2013; Weis et al., 2018); restricting samples to single countries or homogeneous institutional types to limit cross-country definitional differences (Curi et al., 2012; Lafuente and Berbegal-Mirabent, 2019); and augmenting sparse data through direct institutional contact (Vinig and Lips, 2015). These are pragmatic workarounds, not solutions: they preserve internal validity at the cost of generalizability, leaving cross-system comparison structurally out of reach.

Table 2: Common comparability problems and how addressed

Issue	Typical treatment	Citations
Incomplete / inconsistent reporting	Explicitly noted; missing years, partial disclosure, student spin-offs often excluded	Vinig and Lips, 2015; Stankevičienė et al., 2017
Cross-country heterogeneity	Frequently acknowledged; most empirical work remains single-country; few true multi-country datasets	Prónay et al., 2022; Secundo et al., 2017; Fadeyi et al., 2019; Juárez and Montoro-Sánchez, 2019
Non-comparable institutional contexts	Recognized as confounding efficiency scores (field mix, TTO age/mandate, national IP regimes)	Prónay et al., 2022; Lafuente and Berbegal-Mirabent, 2019; Curi et al., 2012; Cartaxo and Godinho, 2017

3.4 Alignment with emerging professional frameworks

ASTP and similar associations have developed harmonized metric frameworks (Barjak et al., 2025), but empirical practice has adopted them only partially, continuing to rely on patent- and license-based indicators.

Some convergence is visible: multi-stage pipelines echo ASTP's value-chain logic (Weis et al., 2018; Lafuente and Berbegal-Mirabent, 2019), and recent European and Japanese TTO work explicitly incorporates intangible dimensions — embeddedness, marketing capabilities — aligning with ASTP's push for holistic assessment (Prónay et al., 2022). What remains absent is direct adoption of ASTP's framework: most studies rely on AUTM definitions or ad-hoc indicator sets (Bradley et al., 2013; de Falani and Torkomian, 2024; Gulbrandsen and Rasmussen, 2012; Sorensen and Chambers, 2008), societal outcomes remain weakly represented, and cross-country harmonization in line with ASTP/EU guidelines is only partially realized (Prónay et al., 2022; Weis et al., 2018).

3.5 Geographical coverage of UTTE studies

The literature is dominated by single-country analyses benchmarked against the United States. National systems studied include Spain (Caldera and Debande, 2010; Lafuente and Berbegal-Mirabent, 2019), Italy (Ramaciotti and Rizzo, 2015), Taiwan (Hsu et al., 2015; Zhang and Zeng, 2025), Indonesia (Sutopo et al., 2019), China (Wang and Zhao, 2022; Li and Fu, 2021), the US (Kim et al., 2008; Anderson et al., 2007; Showalter and Jensen, 2019), South Korea (Lee and Jung, 2021; Hwang et al., 2023), Japan (Modic and Suklan, 2022), Brazil (Leite et al., 2023), Australia (Nugent and Chan, 2023), Canada (Moroz et al., 2008), Turkey (Ustundag et al., 2011; Belgin et al., 2025), South Africa (Gumbi, 2010), and Norway (Gulbrandsen and Rasmussen, 2012), many narrowed to small institutional samples. The one multi-continent exception — a Europe–Japan survey of 18 countries — explicitly notes that single-country work and US-centric benchmarking dominate the field (Prónay et al., 2022; Lee and Jung, 2024; de Falani and Torkomian, 2024). Efficiency benchmarks calibrated to Bayh-Dole-era US institutions are weakly comparable when applied to European contexts.

4 Conclusion

This systematic review of 71 studies reveals a field constrained by four structural limitations: single-country scope that precludes cross-system comparison; supply-side framing that ignores downstream economic value; continued reliance on patent and licensing proxies despite calls for broader impact metrics; and weak measurement of the organizational mechanisms — capabilities, embeddedness, trust — that shape transfer outcomes.

These findings carry a theoretical implication. The dominant paradigm models technology transfer as a unilateral production process, yet Triple Helix theory frames it as reciprocal interaction between university, industry, and government. What the literature measures as efficiency is better understood as a transactional proxy for a translational process oriented toward co-creating value across organizational boundaries. The gap between available metrics and theoretically relevant outcomes suggests the field requires a reconceptualization of what UTTE means in a demand-side perspective.

These findings carry practical implications. TTO managers should recognize that current benchmarking tools may systematically undervalue offices pursuing non-IP channels such as contract research or collaborative development. Policymakers designing performance frameworks — particularly in EU widening countries — should adopt ASTP-type harmonized metrics rather than importing US-centric indicators. University administrators should look beyond patent counts toward broader commercialization outcomes and relational capabilities. We propose three research directions: shifting the unit of analysis to the national innovation ecosystem using European Innovation Scoreboard data to relate input configurations to outcome measures across EU member states; examining innovation-driven enterprises as downstream recipients of transferred knowledge; and qualitative inquiry into the micro-level mechanisms that macro-level analysis cannot capture.

5 Ethics Declaration

This study is based on a systematic review of published peer-reviewed literature and does not involve human subjects, personal data, or primary data collection. No ethical clearance was required.

6 AI Declaration

Two AI tools were used in developing this paper. Consensus AI was used to support the analysis of the reviewed literature and to identify patterns across studies. Claude AI (Anthropic) was used to assist in constructing the analytical framework and methodology, and for text revisions including language editing, reference formatting, and structural refinement. In all cases, the authors retained full responsibility for the intellectual content, interpretation of findings, and argumentation.

References

Anderson, T., Daim, T. and Lavoie, F.F. (2007) "Measuring the efficiency of university technology transfer", *Technovation*, Vol. 27, No. 5, pp 306–318, doi: 10.1016/j.technovation.2006.10.003.

Barjak, F., Bailey, A., Schmidt, T., Lefebvre, F. and Spinardi, L. (2025) *Knowledge and Technology Transfer Metrics Report: Data Report from ASTP National Associations Advisory Committee (NAAC) Working Group on KPI Metrics & Impact*, ASTP, available at: https://astp4kt.eu/resources/ktt-metrics-report-to-boost-harmonisation-across-europe-released.pdf

Barra, C. and Zotti, R. (2018) "The contribution of university, private and public sector resources to Italian regional innovation system (in)efficiency", *Journal of Technology Transfer*, Vol. 43, No. 2, pp 432–457, doi: 10.1007/s10961-016-9539-7.

Belgin, Ö., Avşar, B.A. and Çekiciler, C. (2025) "Evaluating the Performance of Transfer Offices Using MCDM Approaches", *Journal of the Knowledge Economy*, Vol. 16, No. 5, pp 16474–16501, doi: 10.1007/s13132-024-02522-7.

Bigliardi, B., Galati, F., Marolla, G. and Verbano, C. (2015) "Factors affecting technology transfer offices' performance in the Italian food context", *Technology Analysis and Strategic Management*, Vol. 27, No. 4, pp 361–384, doi: 10.1080/09537325.2014.1002464.

Bradley, S., Hayter, C. and Link, A. (2013) "Models and Methods of University Technology Transfer", *Foundations and Trends in Entrepreneurship*, Vol. 9, No. 6, pp 571–650, doi: 10.1561/0300000048.

Caldera, A. and Debande, O. (2010) "Performance of Spanish universities in technology transfer: An empirical analysis", *Research Policy*, Vol. 39, No. 9, pp 1160–1173, doi: 10.1016/j.respol.2010.05.016.

Cartaxo, R.M. and Godinho, M.M. (2017) "How institutional nature and available resources determine the performance of technology transfer offices", *Industry and Innovation*, Vol. 24, No. 7, pp 713–734, doi: 10.1080/13662716.2016.1264068.

Chapple, W., Lockett, A., Siegel, D. and Wright, M. (2005) "Assessing the relative performance of UK university technology transfer offices: parametric and non-parametric evidence", *Research Policy*, Vol. 34, No. 3, pp 369–384, doi: 10.1016/j.respol.2005.01.007.

Chen, K., Wupur, A., Liu, X. and Yang, G. (2024a) "Measuring Multiple-Path Technology Transfer Efficiency in Chinese Universities: A Network DEA-Tobit Approach", *Journal of the Knowledge Economy*, doi: 10.1007/s13132-024-01971-4.

Chen, X., Guan, Z., Yang, G., Pan, H. and Xiong, X. (2024b) "Evaluation of technology transfer performance for Chinese universities based on a dual-game cross-efficiency model", *Socio-Economic Planning Sciences*, Vol. 94, doi: 10.1016/j.seps.2024.101950.

Conti, A. and Gaulé, P. (2011) "Is the US outperforming Europe in university technology licensing? A new perspective on the European Paradox", *Research Policy*, Vol. 40, No. 1, pp 123–135, doi: 10.1016/j.respol.2010.10.007.

Coupet, J. and Ba, Y. (2022) "Benchmarking university technology transfer performance with external research funding: a stochastic frontier analysis", *Journal of Technology Transfer*, Vol. 47, No. 2, pp 605–620, doi: 10.1007/s10961-021-09856-3.

Curi, C., Daraio, C. and Llerena, P. (2012) "University technology transfer: How (in)efficient are French universities?", *Cambridge Journal of Economics*, Vol. 36, No. 3, pp 629–654, doi: 10.1093/cje/bes020.

Curi, C., Daraio, C. and Llerena, P. (2015) "The productivity of French technology transfer offices after government reforms", *Applied Economics*, Vol. 47, No. 28, pp 3008–3019, doi: 10.1080/00036846.2015.1011318.

Djokovic, D. and Souitaris, V. (2008) "Spinouts from academic institutions: a literature review with suggestions for further research", *Journal of Technology Transfer*, Vol. 33, No. 3, pp 225–247, doi: 10.1007/s10961-006-9000-4.

Fadeyi, O., Marešová, P., Štemberková, R., Afolayan, M. and Adeoye, F. (2019) "Perspectives of University-Industry technology transfer in African emerging economies: Evaluating the Nigerian scenario via a data envelopment approach", *Social Sciences*, Vol. 8, No. 10, doi: 10.3390/socsci8100286.

de Falani, S.Y.A. and Torkomian, A.L.V. (2024) "Technology Transfer Offices: a Systematic Review of the Literature and Future Perspective", *Journal of the Knowledge Economy*, Vol. 15, No. 1, pp 4455–4488, doi: 10.1007/s13132-023-01319-4.

Feng, H.-I., Chen, C.-S., Wang, C.-H. and Chiang, H.-C. (2012) "The role of intellectual capital and university technology transfer offices in university-based technology transfer", *Service Industries Journal*, Vol. 32, No. 6, pp 899–917, doi: 10.1080/02642069.2010.545883.

Fernández-Alles, M., Rodríguez-González, T., Villanueva-Flores, M. and Díaz-Fernández, M. (2025) "Impact of the ambidexterity of University Technology Transfer Offices on their performance: The Spanish context", *Technological Forecasting and Social Change*, Vol. 218, doi: 10.1016/j.techfore.2025.124224.

Gulbrandsen, M. and Rasmussen, E. (2012) "The use and development of indicators for the commercialisation of university research in a national support programme", *Technology Analysis & Strategic Management*, Vol. 24, No. 5, pp 481–495, doi: 10.1080/09537325.2012.674670.

Gumbi, S. (2010) "A review of performance standards to monitor, evaluate and assess the impact of technology transfer offices", *South African Journal of Science*, Vol. 106, No. 7, doi: 10.4102/sajs.v106i7/8.323.

Heisey, P.W. and Adelman, S.W. (2011) "Research expenditures, technology transfer activity, and university licensing revenue", *Journal of Technology Transfer*, Vol. 36, No. 1, pp 38–60, doi: 10.1007/s10961-009-9129-z.

Hsu, D.W.L., Shen, Y.-C., Yuan, B.J.C. and Chou, C.J. (2015) "Toward successful commercialization of university technology: Performance drivers of university technology transfer in Taiwan", *Technological Forecasting and Social Change*, Vol. 92, pp 25–39, doi: 10.1016/j.techfore.2014.11.002.

Huang, T., Ken, Y., Wang, W.-C., Wu, C.-H. and Shiu, S.-H. (2011) "Assessing the relative performance of U.S. university technology transfer: Non-parametric evidence", *WSEAS Transactions on Business and Economics*, Vol. 8, No. 3, pp 79–109.

Huang, Y., Zhang, J. and Xu, Y. (2025) "Which configurations promote the high-level university technology transfer? Evidence from China", *Higher Education*, doi: 10.1007/s10734-025-01513-0.

Hülsbeck, M., Lehmann, E.E. and Starnecker, A. (2013) "Performance of technology transfer offices in Germany", *Journal of Technology Transfer*, Vol. 38, No. 3, pp 199–215, doi: 10.1007/s10961-011-9243-6.

Hwang, K., Cho, D. and Shenkoya, T. (2023) "Industry-Academic cooperation in Korea: an analysis of the effect of human resources", *Technology Analysis & Strategic Management*, Vol. 35, No. 3, pp 339–351, doi: 10.1080/09537325.2021.1975034.

Iacobucci, D., Micozzi, A. and Piccaluga, A.M.C. (2021) "An empirical analysis of the relationship between university investments in Technology Transfer Offices and academic spin-offs", *R and D Management*, Vol. 51, No. 1, pp 3–23, doi: 10.1111/radm.12434.

Juárez, J.A.Y. and Montoro-Sánchez, M.Á.M. (2019) "Efficiency and productivity in transfer units of scientific research results in Mexico", *Contaduria y Administracion*, Vol. 64, No. 3, doi: 10.22201/fca.24488410e.2019.1421.

Kim, J., Anderson, T. and Daim, T. (2008) "Assessing university technology transfer: A measure of efficiency patterns", *International Journal of Innovation and Technology Management*, Vol. 5, No. 4, pp 495–526, doi: 10.1142/S0219877008001497.

Kim, Y. (2013) "The ivory tower approach to entrepreneurial linkage: Productivity changes in university technology transfer", *Journal of Technology Transfer*, Vol. 38, No. 2, pp 180–197, doi: 10.1007/s10961-011-9217-8.

Kreiling, L. and Bounfour, A. (2020) "A practice-based maturity model for holistic TTO performance management: development and initial use", *Journal of Technology Transfer*, Vol. 45, No. 6, pp 1718–1747, doi: 10.1007/s10961-019-09756-7.

Kurman, M. (2011) "An index-based measure of university technology transfer", *International Journal of Innovation Science*, Vol. 3, No. 4, pp 167–176, doi: 10.1260/1757-2223.3.4.167.

Lafuente, E. and Berbegal-Mirabent, J. (2019) "Assessing the productivity of technology transfer offices: an analysis of the relevance of aspiration performance and portfolio complexity", *Journal of Technology Transfer*, Vol. 44, No. 3, pp 778–801, doi: 10.1007/s10961-017-9604-x.

Lee, J., Lee, S., Kim, C.-H. and Yoon, J. (2023) "Technology-Transferability Analysis of Universities and Public Research Institutes Using Deep Neural Networks", *IEEE Access*, Vol. 11, pp 135196–135211, doi: 10.1109/ACCESS.2023.3337830.

Lee, K. and Jung, H.J. (2021) "Does TTO capability matter in commercializing university technology? Evidence from longitudinal data in South Korea", *Research Policy*, Vol. 50, No. 1, doi: 10.1016/j.respol.2020.104133.

Lee, K. and Jung, H.J. (2024) "What makes universities build academic spin-offs more successfully? A theory-based triangulation of quantitative studies based on meta-analyses", *Journal of Business Venturing Insights*, Vol. 21, doi: 10.1016/j.jbvi.2024.e00463.

Leite, R.A.S., Walter, C.E., Reis, I.B., de Sousa, P.E.F., De Aragão Gomes, I.M. and Au-Yong-Oliveira, M. (2023) "Towards a model for determining patent revenue odds: An empirical study of technology transfer offices", *Expert Systems*, Vol. 40, No. 5, doi: 10.1111/exsy.13037.

Li, D. and Fu, X. (2021) "The efficiency of university technologtransfer in china", *Asian Economic Papers*, Vol. 20, No. 3, doi: 10.1162/asep_a_00838.

Li, X. and Justin Tan, J. (2020) "Exploring the Role of University-Run Enterprises in Technology Transfer from Chinese Universities", *Management and Organization Review*, Vol. 16, No. 4, pp 907–943, doi: 10.1017/mor.2019.55.

Liu, S., He, M., Zhou, Q., Ma, J. and Zhao, Z. (2026) "Research on the efficiency of technology transfer in Chinese universities from the perspective of supply and demand", *Socio-Economic Planning Sciences*, Vol. 104, doi: 10.1016/j.seps.2025.102382.

Lynskey, M.J. (2008) "The Entrepreneurial University and Spin-Out Firms in the UK: Playing the Numbers Game?", *Industry and Higher Education*, Vol. 22, No. 2, pp 81–98, doi: 10.5367/000000008784139505.

Ma, D., Cai, Z. and Zhu, C. (2022) "Technology transfer efficiency of universities in China: A three-stage framework based on the dynamic network slacks-based measurement model", *Technology in Society*, Vol. 70, doi: 10.1016/j.techsoc.2022.102031.

Modic, D. and Suklan, J. (2022) "Multidimensional experience and performance of highly skilled administrative staff: Evidence from a technology transfer office", *Research Policy*, Vol. 51, No. 10, doi: 10.1016/j.respol.2022.104562.

Moroz, P.W., Hindle, K.G. and Anderson, R.B. (2008) "Commercialisation of new knowledge within universities: Exploring performance disparities", *International Journal of Technology Intelligence and Planning*, Vol. 4, No. 1, pp 4–19, doi: 10.1504/IJTIP.2008.017399.

Nugent, A. and Chan, H.F. (2023) "Outsourcing university research commercialization to a sophisticated technology transfer office: Evidence from Australian universities", *Technovation*, Vol. 125, doi: 10.1016/j.technovation.2023.102762.

Prónay, S., Keszey, T., Buzás, N., Sakai, T. and Inai, K. (2022) "Performance of university technology transfer offices: evidence from Europe and Japan", *International Journal of Productivity and Performance Management*, Vol. 71, No. 4, pp 1343–1364, doi: 10.1108/IJPPM-03-2020-0091.

Pujotomo, D., Ma'aram, A., Isyraf, M.I., Syed Hassan, S.A.H. and Sutopo, W. (2025) "The scenario of accelerating technology commercialization at research university, a system dynamics approach", *Journal of Open Innovation: Technology, Market, and Complexity*, Vol. 11, No. 1, doi: 10.1016/j.joitmc.2025.100497.

Ramaciotti, L. and Rizzo, U. (2015) "The determinants of academic spin-off creation by Italian universities", *R and D Management*, Vol. 45, No. 5, pp 501–514, doi: 10.1111/radm.12105.

Sallan, J. and Lordan, O. (2025) "University technology transfer: leveraging experiential learning and TTO's resources", *Small Business Economics*, Vol. 64, No. 1, pp 221–237, doi: 10.1007/s11187-024-00899-y.

Scanlan, J. (2018) "A capability maturity framework for knowledge transfer", *Industry and Higher Education*, Vol. 32, No. 4, pp 235–244, doi: 10.1177/0950422218770335.

Secundo, G., de Beer, C., Fai, F.M. and Schutte, C. (2019) "Increasing university entrepreneurialism: qualitative insights from the technology transfer office", *Measuring Business Excellence*, Vol. 23, No. 3, pp 253–268, doi: 10.1108/MBE-02-2019-0015.

Secundo, G., de Beer, C., Schutte, C. and Passiante, G. (2017) "Mobilising intellectual capital to improve European universities' competitiveness: The technology transfer offices' role", *Journal of Intellectual Capital*, Vol. 18, No. 3, pp 607–624, doi: 10.1108/JIC-12-2016-0139.

See, K.F., Ma, Z. and Tian, Y. (2023) "Examining the efficiency of regional university technology transfer in China: A mixed-integer generalized data envelopment analysis framework", *Technological Forecasting and Social Change*, Vol. 197, doi: 10.1016/j.techfore.2023.122802.

Showalter, D. and Jensen, R. (2019) "University startup intensity and faculty quality", *Economic Inquiry*, Vol. 57, No. 2, pp 855–875, doi: 10.1111/ecin.12752.

Siegel, D., Waldman, D. and Link, A. (2003) "Assessing the impact of organizational practices on the relative productivity of university technology transfer offices: an exploratory study", *Research Policy*, Vol. 32, No. 1, pp 27–48, doi: 10.1016/S0048-7333(01)00196-2.

Siegel, D.S., Wright, M., Chapple, W. and Lockett, A. (2008) "Assessing the relative performance of university technology transfer in the US and UK: A stochastic distance function approach", *Economics of Innovation and New Technology*, Vol. 17, No. 7, pp 717–729, doi: 10.1080/10438590701785769.

Sorensen, J.A.T. and Chambers, D.A. (2008) "Evaluating academic technology transfer performance by how well access to knowledge is facilitated - Defining an access metric", *Journal of Technology Transfer*, Vol. 33, No. 5, pp 534–547, doi: 10.1007/s10961-007-9038-y.

Stankevičienė, J., Kraujalienė, L. and Vaiciukevičiūtė, A. (2017) "Assessment of technology transfer office performance for value creation in higher education institutions", *Journal of Business Economics and Management*, Vol. 18, No. 6, pp 1063–1081, doi: 10.3846/16111699.2017.1405841.

Sutopo, W., Astuti, R.W. and Suryandari, R.T. (2019) "Accelerating a technology commercialization; with a discussion on the relation between technology transfer efficiency and open innovation", *Journal of Open Innovation: Technology, Market, and Complexity*, Vol. 5, No. 4, doi: 10.3390/joitmc5040095.

Sutopo, W., Khofiyah, N.A., Hisjam, M. and Ma'aram, A. (2022) "Performance Efficiency Measurement Model Development of a Technology Transfer Office (TTO) to Accelerate Technology Commercialization in Universities", *Applied System Innovation*, Vol. 5, No. 1, doi: 10.3390/asi5010021.

Tseng, A.A. and Raudensky, M. (2014a) "Performance evaluations of technology transfer offices of major US research universities", *Journal of Technology Management and Innovation*, Vol. 9, No. 1, pp 93–102, doi: 10.4067/S0718-27242014000100008.

Tseng, A.A. and Raudensky, M. (2014b) "Assessments of technology transfer activities of US universities and associated impact of Bayh–Dole Act", *Scientometrics*, Vol. 101, No. 3, pp 1851–1869, doi: 10.1007/s11192-014-1404-6.

Ustundag, A., Yanik, S. and Kilinç, M.S. (2011) "Evaluating the performance of technology transfer offices", *Journal of Enterprise Information Management*, Vol. 24, No. 4, pp 322–337, doi: 10.1108/17410391111148576.

Villasalero, M. (2014) "University knowledge, open innovation and technological capital in Spanish science parks: Research revealing or technology selling?", *Journal of Intellectual Capital*, Vol. 15, No. 4, pp 479–496, doi: 10.1108/JIC-07-2014-0083.

Vinig, T. and Lips, D. (2015) "Measuring the performance of university technology transfer using meta data approach: the case of Dutch universities", *Journal of Technology Transfer*, Vol. 40, No. 6, pp 1034–1049, doi: 10.1007/s10961-014-9389-0.

Wang, J. and Zhao, M. (2022) "Evaluation of regional university technology transfer performance in China based on super-efficiency SBM", *International Journal of Technology Management*, Vol. 90, No. 1, pp 54–77, doi: 10.1504/IJTM.2022.10048638.

Wang, X. (2025) "How to promote university technology transfer? A configuration analysis based on technology, organization and environment framework", *PLOS ONE*, Vol. 20, No. 3, doi: 10.1371/journal.pone.0318563.

Weis, J., Bashyam, A., Ekchian, G.J., Paisner, K. and Vanderford, N.L. (2018) "Evaluating disparities in the U.S. technology transfer ecosystem to improve bench to business translation.", *Wellcome Open Research*, Vol. 7, doi: 10.12688/f1000research.14210.1.

Xia, Q., Cao, Q. and Tan, M. (2020) "Basic research intensity and diversified performance: the moderating role of government support intensity", *Scientometrics*, Vol. 125, No. 1, pp 577–605, doi: 10.1007/s11192-020-03635-x.

Yang, W. and Chen, W. (2025) "A configurational exploration of how organizational characteristics and multidimensional proximities influence university-industry collaborations", *Journal of Technology Transfer*, doi: 10.1007/s10961-025-10272-0.

Zarea, F., Douglas, E.J., Obschonka, M., Davidsson, P., Audretsch, D.B. and Hutmacher, D.W. (2025) "When the marketplace comes to the research laboratory: technology transfer efficiency of innovation-driven publicly funded research centers", *Journal of Technology Transfer*, Vol. 50, No. 6, pp 2556–2586, doi: 10.1007/s10961-025-10188-9.

Zhang, M. and Zeng, J. (2025) "Which governance structures are conducive to the performance of TTOs? Evidence from Taiwan", *Journal of Technology Transfer*, Vol. 50, No. 1, pp 271–303, doi: 10.1007/s10961-024-10097-3.

A Hybrid RAG Architecture for Knowledge Management in Urban Digital Twins

Giovanni Felle, Antonio Arcadi, Davide Storelli, Valentina Chetta, Rosaria Lovecchio, Elisa De Bellis
Engineering Ingegneria Informatica S.p.A., Italy

Abstract

Urban Digital Twins (UDTs) have emerged as strategic instruments for urban planning and governance, generating unprecedented volumes of heterogeneous knowledge assets — policy documents, municipal acts, citizen reports, and sensor data streams — that overwhelm traditional Knowledge Management (KM) frameworks. Conventional AI-driven systems struggle to deliver accurate, traceable responses on complex, multi-domain queries, exhibiting persistent hallucination and insufficient contextual grounding: benchmarking studies reveal that advanced LLMs achieve only 34% accuracy on complex queries, standard Retrieval-Augmented Generation (RAG) improves this to merely 44%, and state-of-the-art solutions answer only 63% of questions without hallucination. These limitations are particularly critical in regulated public governance contexts where transparency, provenance, and temporal consistency are fundamental requirements. This paper presents a Hybrid Retrieval-Augmented Generation (H-RAG) architecture for knowledge management in UDT environments, integrating vector-based semantic retrieval with Knowledge Graph-grounded reasoning. Following a Design Science Research (DSR) methodology, we describe the end-to-end architecture of a Knowledge Management System (KMS) comprising six interconnected modules: heterogeneous document ingestion, parsing and normalisation, transformer-based embedding generation, incremental Knowledge Graph construction, hybrid multimodal indexing, and context-fused retrieval and generation. A Knowledge Intake Guard (KIG) ensures semantic consistency and deduplication prior to graph ingestion, preventing knowledge base fragmentation as the document corpus evolves. The KMS adopts a microservices-oriented pattern with asynchronous processing (message broker and task queue) and multimodal storage (object store, vector database, graph database, and search engine), enabling scalable ingestion of structured and unstructured municipal document corpora. Unlike batch-oriented graph approaches, the proposed system supports incremental knowledge enrichment without requiring periodic full reconstruction of the knowledge base. An illustrative case study based on a municipal traffic ordinance demonstrates how the architecture supports cross-silo knowledge integration, serving both citizens submitting natural-language queries and technical operators requiring entity-centric relational retrieval. Benchmarking evidence from the literature confirms substantial accuracy improvements in Knowledge Graph-augmented retrieval settings, with a corresponding significant reduction in hallucinations. The proposed KMS is developed within the AVANT IPCEI-CIS project and contributes: (i) a conceptual framework for hybrid knowledge retrieval in dynamic urban environments; (ii) a replicable DSR-grounded design approach applicable to public administration KM; and (iii) architectural guidelines and design principles for integrating governance-oriented KM capabilities into multi-silo UDT contexts.

Keywords: Hybrid RAG (H-RAG), Urban Digital Twins, Knowledge Graph (KG), Urban Knowledge Management, Evidence-Based Governance.

Paper type: Practical Paper

1 Introduction

Urban Digital Twins (UDTs) have emerged as strategic instruments for urban planning and governance, generating unprecedented volumes of heterogeneous knowledge artifacts including policy documents, citizen science data, analytical reports, and real-time sensor datasets (Deren, Wenbo and Zhenfeng, 2021). Managing this dynamic and multi-domain knowledge ecosystem poses significant challenges for traditional KM frameworks, which often struggle to maintain pace with the rapid evolution and sheer volume of unstructured data distributed across administrative silos.

RAG has become a standard architectural pattern for incorporating domain-specific knowledge into AI-powered applications (Gao et al., 2023). However, benchmarking studies reveal significant limitations: advanced LLMs achieve only 34% accuracy on complex queries, while standard RAG improves this to merely 44%, and state-of-the-art solutions answer only 63% of questions without hallucination (Yang et al., 2024). Moreover, recent comparative evaluations of RAG versus GraphRAG confirm that graph-augmented approaches provide qualitatively distinct capabilities for entity-centric and relational queries (Han et al., 2025). These limitations hinder the adoption of AI-driven KM systems in regulated urban governance contexts where transparency, provenance, and temporal consistency are fundamental requirements. In particular, the digitisation of municipal administrative acts — decrees, resolutions, ordinances — presents significant challenges for searchability and

interoperability that standard retrieval paradigms cannot fully address, as these documents create intricate webs of regulatory dependencies and cross-references that demand structured relational reasoning.

While existing literature has explored RAG and GraphRAG separately, few studies have addressed their systematic integration for urban knowledge bases. This paper addresses these limitations by proposing a Hybrid RAG (H-RAG) architecture for UDT environments, developed within the AVANT IPCEI-CIS project. The system targets three primary user profiles: public administrators, citizens, and technical operators requiring decision support grounded in verified, traceable urban knowledge.

The following research question guides this work: To what extent can a hybrid architecture integrating semantic, lexical, and graph-based retrieval address the knowledge fragmentation challenge across administrative silos in Urban Digital Twin environments?

2 Related Work

2.1 Knowledge Management and AI-Augmented Systems

Knowledge Management (KM) is broadly defined as the deliberate, integrated process of identifying, capturing, and leveraging collective organisational knowledge to support effective action and decision-making. KM efforts revolve around four interdependent processes — creation, storage and retrieval, sharing, and application — forming the KM lifecycle (Alavi and Leidner, 2001). The adoption of AI-based approaches has significantly expanded the capacity of KM systems to handle unstructured data and complex retrieval tasks (Jarrahi et al., 2023). In the UDT context, KM challenges are compounded by the multi-domain, temporally layered nature of urban knowledge spanning multiple administrative silos.

2.2 From Information Retrieval to Hybrid Retrieval

The evolution of information retrieval traces from the Vector Space Model and probabilistic BM25 ranking (Robertson and Walker, 1994), through Latent Semantic Indexing (Deerwester et al., 1990), to dense neural retrieval powered by transformer-based embeddings (Devlin et al., 2019; Reimers and Gurevych, 2019; Karpukhin et al., 2020). Hybrid retrieval approaches combining dense vector search with sparse keyword search have demonstrated superior performance over single-paradigm retrieval on heterogeneous document collections (Kiran, 2025), motivating their adoption in the proposed architecture.

2.3 Retrieval-Augmented Generation and Knowledge Graph Augmentation

RAG architectures were introduced to mitigate LLM limitations — hallucination and static knowledge cutoffs — by conditioning generation on dynamically retrieved document passages (Lewis et al., 2020). Recent surveys synthesise the rapidly expanding RAG landscape and its variants (Gao et al., 2023; Fan et al., 2024). Recent research has extended RAG toward semantic compression (Raja and Vats, 2025) and multimodal retrieval for industrial applications (Riedler and Langer, 2024).

Knowledge Graphs (KGs) provide structured, explicit representations of entities and semantic relationships, enabling deterministic graph traversal for multi-hop queries. The integration of KGs with LLM-based retrieval reduces hallucination rates and improves accuracy on complex enterprise knowledge tasks: Agrawal et al. (2024) confirm that KG grounding substantially reduces hallucinations in LLMs, while Sequeda, Allemang and Jacob (2023) report an accuracy increase from 16.7% to 54.2% when KG grounding is combined with LLM-based retrieval. Sarmah et al. (2024) propose HybridRAG as a practical integration pattern combining KG and vector retrieval, and Edge et al. (2024) extend this paradigm to graph-based query-focused summarisation. Neural-symbolic hybrid approaches further combine the interpretability of symbolic graph reasoning with the robustness of neural embeddings (Futia and Vetrò, 2020; Zhang et al., 2021). To the authors' knowledge, no prior work has specifically addressed the design of an H-RAG architecture tailored to the multi-silo UDT environment and the governance requirements of public administration KM.

3 Methodology

This research follows the Design Science Research (DSR) methodology as formalised by Hevner et al. (2004) and Peffers et al. (2007). DSR is particularly suited to the development and evaluation of innovative IT artifacts in contexts requiring both technical rigour and practical relevance. The methodology proceeds through five iterative phases: problem identification, objectives definition, design and development, demonstration, and evaluation.

Three core design requirements, derived from analysis of the AVANT project's governance and operational context, guide the artifact's design:

- R1: Support for heterogeneous, incrementally ingested document types without requiring periodic full knowledge base reconstruction.
- R2: Transparent, provenance-anchored responses for regulated decision-making, ensuring every generated answer is traceable to specific source documents.
- R3: Cross-silo semantic integration across urban administrative domains, enabling multi-hop reasoning over entities and relationships distributed across multiple document collections.

The system is currently in a prototypical phase deployed in a controlled test environment, with ingestion and end-to-end pipeline validation underway on a corpus of municipal administrative acts.

The demonstration phase, instantiated through the case study presented in Section 5, provides evidence of the architecture's applicability to realistic urban governance scenarios. Full empirical evaluation is planned as part of the AVANT First Industrial Deployment (FID) phase, scheduled for July 2026.

4 Proposed KMS Architecture

4.1 Architecture Overview

The proposed KMS implements a knowledge management infrastructure designed for large-scale municipal document ingestion, semantic processing, and hybrid retrieval in Urban Digital Twin environments. The platform adopts a microservices-oriented pattern with asynchronous processing and a multimodal storage strategy to support scalability and data integrity. At this stage, the implementation prioritises architectural integration and end-to-end data-flow validation over performance optimisation.

The system is organised into six functionally distinct layers, each responsible for a stage of the document lifecycle: ingestion, parsing, vectorisation, indexing, graph construction, and information retrieval.

The three design requirements introduced in Section 3 are addressed as follows: R1 (incremental ingestion without full reconstruction) is realised through the asynchronous pipeline and the Knowledge Intake Guard; R2 (provenance-anchored responses) is enforced through persistent metadata tagging at ingestion and CITES relationships in the Knowledge Graph; R3 (cross-silo semantic integration) is delivered by the hybrid retrieval layer combining vector, lexical, and graph-based search.

Figure 1 illustrates the end-to-end architecture described in the following subsections.

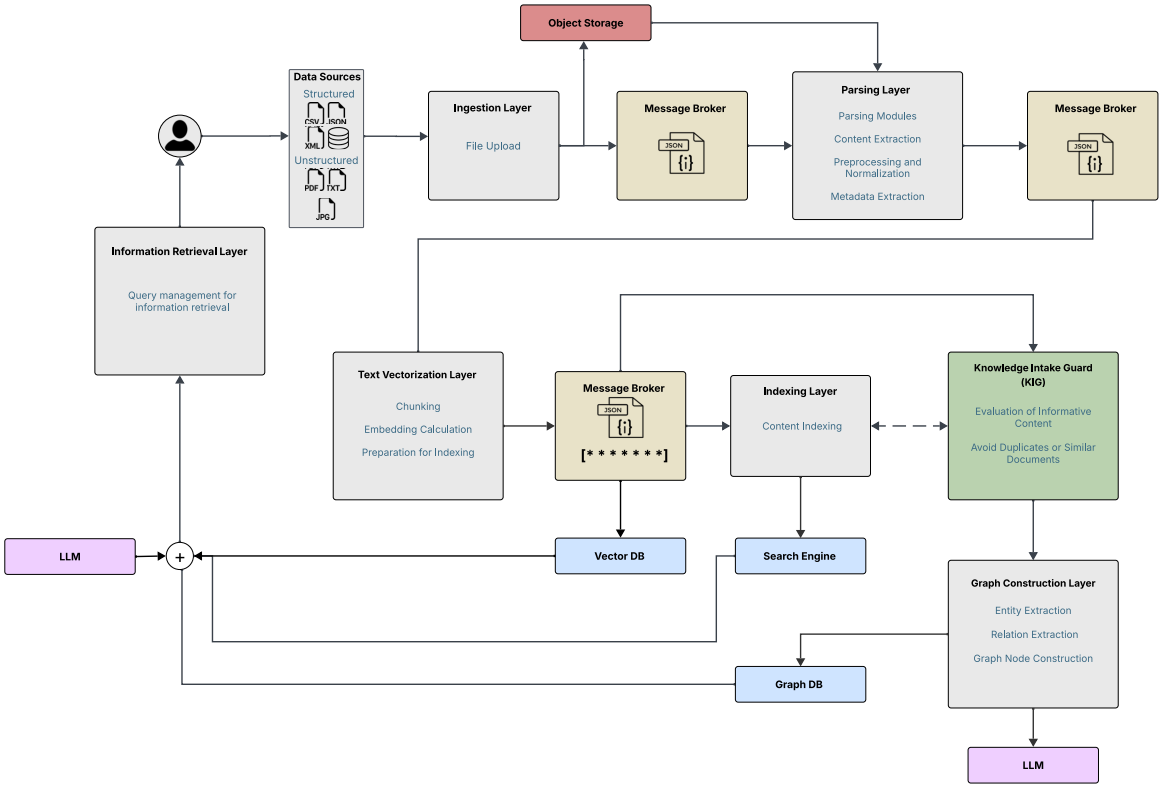


Figure 1: The KMS Architecture

4.2 Technical Description of the KMS Architecture

The workflow originates at the frontend interface, where user sessions and security are managed by a central Identity and Access Management (IAM) layer based on standard authentication and authorisation protocols. This layer enforces role-based access control policies consistently across services.

Data enters through the Ingestion Layer, which exposes a high-throughput API gateway for both structured (CSV, JSON) and unstructured (PDF, TXT) sources. The ingestion service validates schemas, captures core metadata and relationships, and persists raw assets in an S3-compatible object store that acts as the system's source of truth.

To avoid blocking request/response latency, the KMS adopts a producer-consumer pattern: the ingestion service publishes events to a distributed, fault-tolerant message broker, and downstream workers consume and execute pipeline stages asynchronously.

The refined data then moves to the Text Vectorisation Layer, where text is partitioned into optimised chunks before being transformed into high-dimensional vectors.

To handle large documents, the engine implements a sliding window chunking strategy.

The embedding is calculated using a transformer-based model optimised for semantic similarity tasks, producing high-dimensional dense vector representations. This model is specifically designed to capture contextual meaning across heterogeneous textual sources.

Before embedding, chunks undergo light normalisation (whitespace/encoding) to improve consistency for semantic retrieval.

These vectors are indexed in a vector database optimised for ANN search that leverages Hierarchical Navigable Small World (HNSW) graphs to enable efficient approximate nearest neighbour search. The collection is configured with a cosine distance metric consistent with the model's embedding dimensionality, supporting compatibility with semantic similarity computations. Each vector is stored alongside a structured payload containing document metadata (e.g., source, type, creation date), enabling metadata-aware retrieval and post-filtering.

Parallel to the vector indexing, the system populates a Knowledge Graph that extends the retrieval pipeline with structured relational capabilities through a "Text-to-Triple" architecture based on semantic extraction (via an LLM and prompt engineering) and graph-based persistence. Each chunk is processed by an extractor that identifies entities and relations based on a specific Urban Digital Twin ontology.

Extracted triples undergo a strict validation phase. This process cross-references the initial output with the allowed ontology, performing:

The final stage automates the insertion into a graph database.

- *(Document)-[:CITES]->(Entity)*
- *(Entity)-[:RELATED {type: "..."}]->(Entity)*

The presented knowledge graph is an experimental component designed for future GraphRAG capabilities. This process is also governed by the Knowledge Intake Guard, which evaluates information density and utilises deduplication logic to maintain the integrity of the graph. Parallel to vector indexing, the system executes a dedicated Search Engine Layer, where the content is indexed for traditional full-text search and keyword-based retrieval.

To support keyword-based retrieval and structured filtering, the Search Engine relies on a custom index schema designed to align with the document processing pipeline. The index mapping defines both textual and structured fields, enabling hybrid search and metadata-aware querying.

Each document is uniquely identified and enriched with contextual attributes, including source, file path, and ingestion metadata. The full textual content is indexed for traditional inverted-index search, while additional fields support filtering, provenance tracking, and chunk-level granularity.

In particular, the system adopts the following index structure:

- Document Identification: id (keyword) ensures unique document referencing.
- Content Representation: content (text) stores the processed textual chunks, while filename and content_type provide descriptive context.
- Data Provenance and Filtering: Fields such as source, path, and created_at enable traceability and temporal queries.
- Chunk-Level Processing: chunk_id allows fine-grained indexing aligned with the embedding pipeline.
- Metadata Enrichment: metadata (object) stores flexible, schema-less attributes extracted during ingestion.
- Vector Representation: The embedding field is defined as a knn_vector, enabling integration with approximate nearest neighbour (ANN) search algorithms.

- Integrity Control: A hash field is used to detect duplicate content and helps ensure consistency across ingestion cycles.

This schema enables the search engine to operate not only as a full-text engine but also as a complementary component in the hybrid retrieval strategy, supporting filtering, ranking, and alignment with vector-based results.

The final stage of the architecture is the Information Retrieval Layer, which executes a hybrid search strategy. Upon receiving a user query, this layer orchestrates synchronised retrieval across the Search Engine (keyword precision), the Vector DB (semantic similarity), and the Graph DB (relational context). Although vector fields can also be hosted in the search engine for moderate-scale workloads, the architecture includes a dedicated vector database to provide improved performance and operational flexibility for high-dimensional embeddings; this also enables future comparative evaluation between alternative deployment options. During query time, the system encodes the user input into the same vector space and performs similarity search with tuned HNSW parameters to balance recall and latency.

To support document-level relevance, retrieved chunks are aggregated and deduplicated by their originating document, retaining the highest-scoring segment per document as a primary signal while leaving room for future extensions (e.g., multi-segment aggregation via mean pooling or reciprocal rank fusion).

To combine heterogeneous retrieval signals, the system adopts a late-fusion strategy based on score normalisation and weighted aggregation. Similarity scores from semantic (vector-based) and keyword (BM25-based) retrieval are independently normalised (max-score) and then merged through a weighted linear combination, enabling fine-grained control over the ranking process. Overall, the approach balances contextual understanding with lexical precision to support both semantically rich questions and exact-match search scenarios.

Results are dynamically filtered according to the user profile before reaching the Chatbot Layer. In this final phase, the system employs a LLM for high-precision re-ranking, selecting the most relevant document segments to generate a grounded, natural language response that cites its sources directly, providing a seamless and verifiable end-user experience.

5 Illustrative Case Study: Processing A Municipal Traffic Ordinance

This section operationalises the architectural components described in Section 4, illustrating how ingestion, hybrid retrieval, and graph augmentation interact in a realistic scenario drawn from the municipal administration domain. The reference document is Traffic Ordinance No. 118/2026 (Municipality of Fictional City), a synthetic but realistic municipal ordinance, constructed to reflect the structure and language of typical Italian administrative acts, and used here solely for illustrative purposes. It regulates the temporary closure of Via Roma for urban regeneration works. Three distinct interaction scenarios — corresponding to an administrator, a citizen, and a traffic engineer — are used to exercise complementary subsystems of the architecture and to ground the design choices in concrete operational outcomes.

5.1 Document Ingestion and Processing

An administrator with the *document_manager* role uploads the ordinance as a PDF through the frontend interface. The ingestion layer validates the document structure and persists the raw file in the object store before dispatching an asynchronous processing event to the message broker. The subsequent pipeline executes without further human input.

The ordinance exhibits structural characteristics typical of Italian administrative acts: a legal preamble citing antecedent resolutions, a normative body organised as an ordered list of provisions, and a closing monitoring clause. In this case study, chunking follows these structural boundaries; when boundaries are not reliably detected or when cross-paragraph context must be preserved, the system applies a sliding-window overlap strategy. This produces five semantically coherent segments:

- *Chunk A*: preamble and legal references (Resolution No. 42/2026, public safety rationale)
- *Chunk B*: full closure of Via Roma between Piazza Verdi and Via Milano
- *Chunk C*: mandatory traffic deviations (Via Milano eastbound, Via Garibaldi westbound)
- *Chunk D*: signage, barriers, and access exceptions (residents, emergency vehicles, authorised providers)
- *Chunk E*: enforcement period (5 May – 20 June 2026) and daily traffic monitoring provision

Each chunk is embedded and indexed in the Vector DB and Search Engine alongside its metadata payload (source: "ordinance_118_2026", content_type: "traffic_regulation", created_at: "2026-04-15"). In parallel, the Knowledge Graph component extracts the following representative triples via the Text-to-Triple pipeline:

Table 1: Sample extracted triples from the Text-to-Triple pipeline.

Extracted Triples		
Subject	Relation	Object
Ordinance_118_2026	Restricts	Via_roma
Via_roma	Located_in	Fictional_city
Via_roma	Affects	Vehicular_traffic
Resolution_42_2026	Authorises	Ordinance_118_2026
Via_milano	Serves_as	Eastbound_deviation
Via_garibaldi	Serves_as	Westbound_deviation

All entity nodes are linked to the source document via a CITES provenance edge, ensuring full traceability from graph traversal results back to the originating ordinance.

5.2 Scenario 1 — Semantic Retrieval (Citizen Query)

A citizen authenticated with the *citizen* role submits the following query:

"Can I drive through Via Roma next month?"

The query contains no administrative terminology, no reference to the ordinance identifier, and no explicit mention of a closure or regulation. The retrieval layer encodes the query into the shared vector space and performs ANN search. The resulting vector is geometrically proximate to Chunk B and Chunk C, as the embedding model captures the underlying intent — movement restriction on a named road — independently of lexical overlap with the indexed text. The BM25 component contributes results anchored to the surface term "Via Roma", partially overlapping with the semantic signal. After late fusion and per-document deduplication, Chunk B is ranked first, followed by Chunk C and Chunk E.

The IAM filter confirms that the ordinance falls within the access scope of the *citizen* role. The chatbot layer generates the following grounded response:

"Via Roma is closed to vehicular traffic between Piazza Verdi and Via Milano until 20 June 2026, due to urban regeneration works (Ordinance No. 118/2026). Eastbound traffic is diverted via Via Milano, westbound via Via Garibaldi. Residents and emergency vehicles retain controlled access."

This scenario illustrates how semantic retrieval enables effective resolution of natural-language queries over regulatory content even in the absence of lexical overlap with the source document—a capability that purely keyword-based retrieval would not reliably provide in this instance.

5.3 Scenario 2 — Relational Retrieval (Technical Query)

A traffic engineer with the *technical_staff* role submits a structured query:

"Which ordinances issued in 2026 impose restrictions on Via Roma?"

This query favours keyword and structured retrieval: BM25 matches on "Via Roma" and "ordinances", while metadata filtering on *created_at* and *content_type* narrows the candidate set. However, neither vector search nor full-text indexing can resolve this relational dimension with sufficient precision, as the association between via_roma and its restricting ordinance is not co-located in any single chunk. By contrast, the Knowledge Graph provides explicit entity grounding and deterministic relation traversal, enabling the system to answer "which ordinances restrict Via Roma" even when evidence is distributed across chunks or implicit in cross-document references.

The Knowledge Graph resolves this directly: starting from the node *via_roma*, the system traverses the restricts relation in reverse and returns *ordinance_118_2026* as the source entity, enriched with its provenance metadata. In a realistic multi-document setting, the same traversal would return a set of ordinances linked to the entity *via_roma*, enabling comparative analysis across multiple regulatory acts. This result is incorporated into the fusion stage alongside the keyword and semantic signals, producing a ranked response that identifies the ordinance, its enforcement period, and its principal provisions.

This scenario highlights how graph augmentation provides qualitatively distinct retrieval capabilities for entity-centric and relationship-oriented queries—particularly in cases involving implicit or cross-chunk relational dependencies that neither keyword search nor vector similarity can resolve independently.

5.4 Observations

The ordinance use case highlights three operational properties of the KMS that are particularly significant in the public administration domain.

The case study suggests that semantic retrieval extends the effective recall of the system beyond what lexical methods alone can achieve: citizen queries formulated in natural, non-technical language are correctly resolved against regulatory documents that share no significant vocabulary with the query surface form. At the same time, BM25 remains valuable for queries with precise terminological anchors, motivating the hybrid fusion strategy over either paradigm in isolation.

The Knowledge Graph is not merely complementary to vector retrieval but provides structurally distinct query capabilities. Relational queries that require traversing implicit associations between entities — such as identifying all regulations affecting a specific road — are not reliably addressed by semantic similarity or inverted-index lookup, regardless of chunking granularity or embedding quality. The graph layer addresses precisely this class of query.

The provenance mechanism — source citation in the chatbot response and CITES edges in the graph — ensures that every system output remains traceable to its originating document. In a public-sector context, this traceability is a necessary condition for institutional trust and accountability, independent of retrieval effectiveness.

Although the scenario uses a single ordinance, the architecture scales to multi-document corpora, enabling aggregation and comparison across regulatory sources.

6 Discussion and Conclusion

This paper presents a H-RAG architecture for knowledge management in Urban Digital Twin environments, investigating to what extent a hybrid integration of semantic, lexical, and graph-based retrieval can mitigate knowledge fragmentation across administrative silos. Following a DSR methodology, the proposed KMS is designed around three core requirements—incremental ingestion (R1), provenance-anchored responses (R2), and cross-silo semantic integration (R3)—and demonstrated through an illustrative case study based on a municipal traffic ordinance. The research offers three main contributions: (i) a conceptual framework for hybrid knowledge retrieval in dynamic urban environments; (ii) a replicable DSR-based design approach for public administration; (iii) architectural guidelines for integrating governance-oriented KM into UDT contexts.

The case study illustrates how hybrid retrieval can extend recall beyond what lexical methods alone achieve for citizen-facing natural-language queries, while Knowledge Graph augmentation enables structurally distinct capabilities for entity-centric and relational queries that neither semantic similarity nor keyword matching can resolve independently.

From a scientific perspective, the paper advances understanding of how structured and unstructured knowledge can be jointly managed under governance constraints, extending the H-RAG literature to a domain—municipal governance—that has received limited attention in existing research.

Expected outcomes and practical impact are targeted at three stakeholder groups:

- Public administrators and urban data managers → faster, evidence-based access to regulatory knowledge with end-to-end provenance traceability (answer-to-source), supporting accountable decision-making.
- Citizens → natural-language access to complex administrative information without requiring domain expertise, improving accessibility of public services.
- IT architects and platform owners → a modular, microservices-oriented reference architecture for integrating hybrid retrieval and governance-oriented KM capabilities into Digital Twin platforms.

The system is currently deployed as an early-stage prototype, with the primary objective of validating architectural feasibility, interoperability, and end-to-end data flow consistency rather than optimising performance. Evaluation is framed as a progressive, multi-stage process, given the non-stationary ingestion corpus, the heterogeneous retrieval paradigms, and the structural variability of municipal documents.

Accordingly, future work will pursue a structured evaluation roadmap covering retrieval effectiveness (precision, recall, nDCG) across lexical, semantic and hybrid configurations, ablation studies on individual retrieval layers, validation of graph extraction against curated ground-truth subsets, latency and scalability analysis under increasing ingestion and query loads, and end-to-end testing in realistic user-query scenarios.

This staged strategy supports the incremental development of a robust infrastructure for structuring municipal knowledge. Overall, the proposed architecture represents a foundational step toward knowledge-driven access to complex administrative information systems.

7 Acknowledgment

This work is framed within the Project “AVANT” — Project no. IPCEI-CL_0000005 — Application protocol no. 108421 of 14/05/2024 — CUP B89J24002920005 — Grant decree no. 1322 of August 8, 2024 — financed by the European Union — NextGenerationEU (IPCEI Funding).

8 Ethics Declaration

This research did not require ethical clearance, as it does not involve the collection or processing of personal data from human participants. The municipal administrative document corpora referenced in the illustrative case study are synthetic and constructed solely for demonstration purposes.

9 AI Declaration

AI tools (specifically, Claude by Anthropic) were used to support the drafting, structural organisation, and language editing of portions of this paper. All technical content, architectural decisions, scientific claims, and bibliographic references were developed, verified, and validated by the authors. The authors take full responsibility for the accuracy of all content.

References

Agrawal, G., Kumarage, T., Alghamdi, Z. and Liu, H. (2024) ‘Can knowledge graphs reduce hallucinations in LLMs?: A survey’, in Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL 2024), pp. 3947–3960. https://doi.org/10.18653/v1/2024.naacl-long.219.

Alavi, M. and Leidner, D.E. (2001) “Review: Knowledge management and knowledge management systems: conceptual foundations and research issues”, MIS Quarterly, 25(1), pp. 107–136. https://doi.org/10.2307/3250961.

Deerwester, S., Dumais, S.T., Furnas, G.W., Landauer, T.K. and Harshman, R. (1990) “Indexing by latent semantic analysis”, Journal of the American Society for Information Science, 41(6), pp. 391–407. https://doi.org/10.1002/(SICI)1097-4571(199009)41:6<391:AID-ASI1>3.0.CO;2-9.

Deren, L., Wenbo, Y. and Zhenfeng, S. (2021) ‘Smart city based on digital twins’, Computational Urban Science, 1(1), Article 4. https://doi.org/10.1007/s43762-021-00005-y.

Devlin, J., Chang, M.-W., Lee, K. and Toutanova, K. (2019) “BERT: Pre-training of deep bidirectional transformers for language understanding”, in Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT), pp. 4171–4186. https://doi.org/10.18653/v1/N19-1423.

Edge, D., Trinh, H., Cheng, N., Bradley, J., Chao, A., Mody, A., Truitt, S., Metropolitansky, D., Ness, R.O. and Larson, J. (2024) ‘From local to global: A GraphRAG approach to query-focused summarisation’, arXiv. https://doi.org/10.48550/arXiv.2404.16130.

Fan, W., Ding, Y., Ning, L., Wang, S., Li, H., Yin, D., Chua, T.-S. and Li, Q. (2024) ‘A survey on RAG meets LLMs: Towards retrieval-augmented large language models’, arXiv. https://doi.org/10.48550/arXiv.2405.06211.

Futia, G. and Vetrò, A. (2020) “On the Integration of Knowledge Graphs into Deep Learning Models for a More Comprehensible AI—Three Challenges for Future Research”, Information, 11(2), 122. https://doi.org/10.3390/info11020122.

Gao, Y., Xiong, Y., Gao, X., Jia, K., Pan, J., Bi, Y., Dai, Y., Sun, J., Wang, M. and Wang, H. (2023) “Retrieval-Augmented Generation for Large Language Models: A Survey”, arXiv. https://doi.org/10.48550/arXiv.2312.10997.

Han, H., Ma, L., Wang, Y., Shomer, H., Lei, Y., Qi, Z., Guo, K., Hua, Z., Long, B., Liu, H., Aggarwal, C.C. and Tang, J. (2025) ‘RAG vs. GraphRAG: A systematic evaluation and key insights’, arXiv. https://doi.org/10.48550/arXiv.2502.11371.

Hevner, A.R., March, S.T., Park, J. and Ram, S. (2004) “Design science in information systems research”, MIS Quarterly, 28(1), pp. 75–105. https://doi.org/10.2307/25148625.

Jarrahi, M.H., Askay, D., Eshraghi, A. and Smith, P.G. (2023) “Artificial intelligence and knowledge management: A partnership between human and AI”, Business Horizons, 66(1), pp. 87–99. https://doi.org/10.1016/j.bushor.2022.03.002.

Karpukhin, V., Oguz, B., Min, S., Lewis, P., Wu, L., Edunov, S., Chen, D. and Yih, W.-T. (2020) “Dense passage retrieval for open-domain question answering”, in Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP), pp. 6769–6781. https://doi.org/10.18653/v1/2020.emnlp-main.550.

Kiran, S. (2025) “Hybrid Retrieval-Augmented Generation (RAG) Systems with Embedding Vector Databases”, International Journal of Scientific Research in Computer Science, Engineering and Information Technology, 11(2), pp. 2694–2702. https://doi.org/10.32628/CSEIT25112702.

Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W.-T., Rocktäschel, T., Riedel, S. and Kiela, D. (2020) “Retrieval-augmented generation for knowledge-intensive NLP tasks”, in Advances in Neural Information Processing Systems, 33, pp. 9459–9474. https://doi.org/10.48550/arXiv.2005.11401.

Peffers, K., Tuunanen, T., Rothenberger, M.A. and Chatterjee, S. (2007) “A design science research methodology for information systems research”, Journal of Management Information Systems, 24(3), pp. 45–77. https://doi.org/10.2753/MIS0742-1222240302.

Raja, R. and Vats, A. (2025) “Beyond Nearest Neighbors: Semantic Compression and Graph-Augmented Retrieval for Enhanced Vector Search”, arXiv. https://doi.org/10.48550/arXiv.2507.19715.

Reimers, N. and Gurevych, I. (2019) “Sentence-BERT: Sentence embeddings using Siamese BERT-networks”, in Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), pp. 3982–3992. https://doi.org/10.18653/v1/D19-1410.

Riedler, M. and Langer, S. (2024) “Beyond Text: Optimizing RAG with Multimodal Inputs for Industrial Applications”, arXiv. https://doi.org/10.48550/arXiv.2410.21943.

Robertson, S.E. and Walker, S. (1994) “Some simple effective approximations to the 2–Poisson model for probabilistic weighted retrieval”, in Croft, W.B. and van Rijsbergen, C.J. (eds.) Proceedings of the 17th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR ’94). London: Springer, pp. 345–354. https://doi.org/10.1007/978-1-4471-2099-5_24.

Sarmah, B., Hall, B., Rao, R., Patel, S., Pasquali, S. and Mehta, D. (2024) “HybridRAG: Integrating Knowledge Graphs and Vector Retrieval Augmented Generation for Efficient Information Extraction”, arXiv. https://doi.org/10.48550/arXiv.2408.04948.

Sequeda, J., Allemang, D. and Jacob, B. (2023) “A benchmark to understand the role of knowledge graphs on large language model’s accuracy for question answering on enterprise SQL databases”, arXiv. https://doi.org/10.48550/arXiv.2311.07509.

Yang, X., Sun, K., Xin, H., Sun, Y., Bhalla, N., Chen, X., Choudhary, S., Gui, R.D., Jiang, Z.W., Jiang, Z., Kong, L., Moran, B., Wang, J., Xu, Y.E., Yan, A., Yang, C., Yuan, E., Zha, H., Tang, N., Chen, L., Scheffer, N., Liu, Y., Shah, N., Wanga, R., Kumar, A., Yih, W.-T. and Dong, X.L. (2024) “CRAG – Comprehensive RAG Benchmark”, arXiv. https://doi.org/10.48550/arXiv.2406.04744.

Zhang, J., Chen, B., Zhang, L., Ke, X. and Ding, H. (2021) “Neural, symbolic and neural-symbolic reasoning on knowledge graphs”, AI Open, 2, pp. 14–35. https://doi.org/10.1016/j.aiopen.2021.03.001.

A Knowledge-Based Life-Cycle Framework for Sustainable Restoration Decision-Making

Stefano Abbate, Anna Scotto di Santolo
Pegaso University, Naples, Italy

Abstract

Cultural heritage represents a key driver of social identity, cultural continuity, and local economic development, yet it is increasingly exposed to environmental, anthropogenic, and climate-related threats. In this context, restoration and conservation interventions require selecting materials and techniques that ensure technical performance while addressing broader sustainability objectives. However, decision-making in cultural heritage restoration is often based on fragmented criteria, with limited integration of environmental, economic, and social dimensions across the entire life cycle. This study addresses this gap by developing a multi-criteria life-cycle framework to support the selection and prioritisation of sustainable restoration techniques in cultural heritage. The research adopts a systematic literature review (SLR) methodology to identify and analyse contributions on Life Cycle Assessment (LCA), Life Cycle Costing (LCC), and Social Life Cycle Assessment (S-LCA) within the heritage conservation domain. A final sample of 45 peer-reviewed articles was examined through descriptive and content analyses to extract relevant indicators, methodological approaches, and decision variables. The output is a structured multi-criteria life-cycle framework that integrates sustainability indicators across environmental, economic, and social dimensions, including, for example, Global Warming Potential (GWP), natural resource depletion, water footprint, equipment and energy costs, health and safety, employment stability, and community well-being. This set of indicators constitutes a coherent, multidimensional taxonomy that underpins a flexible, replicable decision-support framework, enabling transparent, systematic comparison of alternative restoration strategies. The study offers relevant implications for both practice and policy, supporting more informed resource allocation and the development of sustainability-oriented guidelines for cultural heritage conservation. The proposed framework is explicitly designed to provide a solid knowledge base for future empirical applications, making it an operational tool for decision-makers. By integrating multi-criteria decision-making techniques, it will be possible to weight the heterogeneous criteria identified and prioritise restoration interventions under conditions of uncertainty. Future developments should also focus on empirical validation through case studies and on incorporating stakeholder perspectives to enhance the framework's applicability and robustness.

Keywords: cultural heritage restoration, life-cycle thinking, sustainability

Paper type: Academic Research Paper

1 Background

Cultural heritage represents a fundamental component of collective identity and regional development, embodying historical, social, and symbolic values that extend beyond its material dimension (Vadrucci, 2025). Archaeological sites contribute to cultural continuity, foster social cohesion, boost tourism, and support local economies (Chen and Tan, 2025). However, these heritage assets face increasing threats, including material degradation, climate change-related hazards, urban development pressures, and anthropogenic impacts. Restoration and consolidation efforts in cultural heritage increasingly rely on a wide range of materials and techniques, including traditional methods, innovative consolidants, and advanced composite or nano-engineered formulations. Recently, significant research has focused on enhancing the mechanical performance, durability, and compatibility of these materials with historic substrates, aiming to improve structural stability while preserving the authenticity of heritage assets (Dal Pozzo et al., 2024). Preventive conservation methods and low-impact treatments are also gaining prominence, as they can reduce the frequency and invasiveness of interventions throughout the entire life cycle of cultural heritage sites (Menegaldo et al., 2023). As a result, the growing availability of alternative materials and restoration techniques has introduced new challenges for decision-makers. Solutions that perform well technically or mechanically may differ substantially in terms of environmental impact, long-term economic effects, or social acceptance. For instance, innovative consolidants might reduce maintenance needs but require energy-intensive manufacturing, whereas traditional materials may promote cultural harmony but result in higher long-term costs or environmental impacts. Despite this complexity, selecting restoration strategies often still relies on isolated performance criteria, with limited consideration of life-cycle environmental, economic, and social impacts. Balancing these often-competing goals is one of the most pressing challenges in contemporary cultural heritage restoration (Angrisano et al., 2024).

By employing a systematic literature review, this study aims to develop a structured, replicable, knowledge-driven evaluation system to support decision-making for the restoration and consolidation of cultural heritage assets. Specifically, it seeks to build a multidimensional assessment framework that serves as a basis for comparing different restoration techniques, materials, and intervention strategies, incorporating environmental, economic, and social sustainability perspectives. The system is based on life-cycle thinking and integrates indicators from Life Cycle Assessment (LCA), Life Cycle Costing (LCC), and Social Life Cycle Assessment (S-LCA) to capture the full range of impacts associated with heritage restoration over time. Consequently, this study introduces a useful knowledge-based tool for practitioners, researchers, and policymakers, thereby contributing to more sustainable, transparent, and well-informed conservation strategies in cultural heritage restoration.

2 Methodology

This study employs a systematic literature review (SLR) structured according to established frameworks for evidence synthesis in management and sustainability research (Cerchione and Esposito, 2016; Greenhalgh and Peacock, 2005). The SLR methodology ensures transparency, reproducibility, and methodological rigour in documenting the search, screening, and selection procedures (Centobelli et al., 2017). The review process was organised into two main phases. The initial phase focused on identifying, gathering, and selecting scientific and technical literature related to sustainability assessment in cultural heritage conservation. The second phase involved a descriptive and content analysis of the final sample of articles to identify and incorporate life-cycle–based sustainability indicators related to maintenance and restoration of cultural heritage assets. These indicators were then integrated into a single decision-support tool designed to prioritise interventions.

2.1 Data collection and selection

The sample of articles was retrieved from the Scopus database, chosen for its extensive coverage of engineering, environmental assessment, and sustainability research (Abbate et al., 2023). Peer-reviewed articles and reviews formed the main evidence base. We obtained an initial sample of 71 documents. Subsequently, following PRISMA recommendations (Page et al., 2021), a preliminary title and abstract screening was conducted. Notably, only documents demonstrating sufficient clarity regarding the methodological framework, relevance to the cultural heritage sector, and alignment with life-cycle–based sustainability assessment principles were retained in the review. During this phase, 30 studies were excluded. The remaining 41 papers were analysed in full text, and six were excluded because they employed non-sustainability-focused approaches. Therefore, 35 papers were considered eligible for inclusion in the final sample. In addition to these core studies, a snowballing approach was used (Greenhalgh and Peacock, 2005). This resulted in the identification of 10 additional relevant publications, primarily on multi-criteria decision-making methods and case studies of life-cycle sustainability assessments in cultural heritage settings. As a result, the final sample consists of 45 articles, as shown in Figure 1.

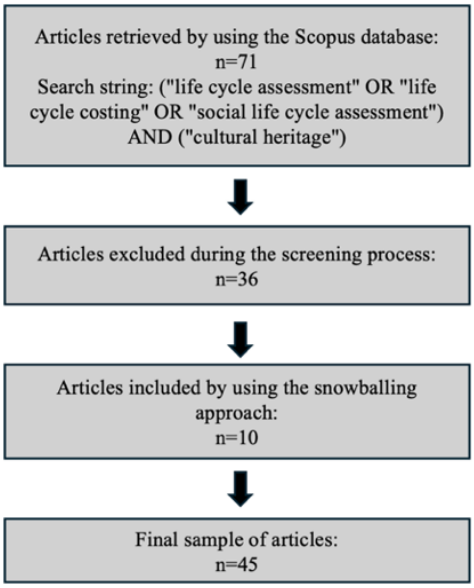


Figure 1. Data collection and selection process

3. Results

3.1 Descriptive statistics

The growing number of publications over time indicates that sustainability in heritage conservation has become an increasingly important research focus. This trend is driven by the global shift towards climate-oriented strategies, efficient use of resources, and the demand for more sustainable conservation approaches. Notably, the increasing application of LCA, LCC, and more recently, SLCA in conservation studies shows a shift from merely technical restoration techniques to broader assessment frameworks that support sustainable decision-making.

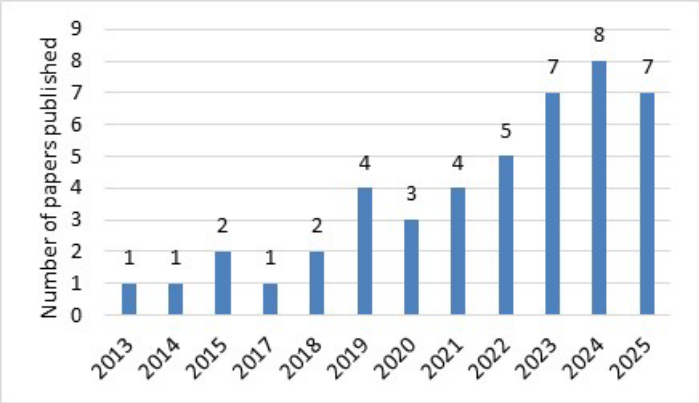


Figure 2. Papers published over time

The analysis of the methodological approaches in the selected publications reveals a predominant use of environmental assessment techniques. Notably, 52% of the papers employ LCA as their primary evaluation tool, confirming its position as the most frequently used sustainability assessment method in the cultural heritage sector. Only 3% of publications combine LCA and LCC, while another 3% integrate LCA with SLCA. Notably, only 7% employ a comprehensive approach that includes LCA, LCC, and SLCA, indicating that a full multidimensional sustainability assessment remains limited and not yet widely adopted. In addition, LCC is used alone in only 3% of the contributions, indicating that economic sustainability is often underrepresented compared to environmental issues. Likewise, only 3% of the papers use SLCA as a standalone method, with the same percentage comprising conceptual studies, which emphasises the necessity for more empirical frameworks. Furthermore, 3% of the publications employ mixed-method approaches, incorporating multi-criteria decision-making (MCDM) techniques; 10% are based on simulation models; and 13% present systematic or narrative literature reviews. This distribution indicates that studies on the topic remain mainly focused on environmental aspects. Economic and social aspects are less commonly examined, and when addressed, they are rarely integrated into a unified assessment framework. The limited adoption of LCC and SLCA, whether used independently or in combination with LCA, highlights a fragmented approach to assessing sustainability in restoration decision-making.

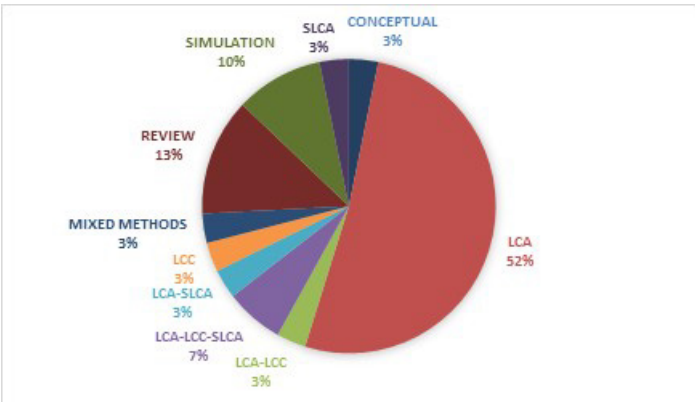


Figure 3. Distribution of papers by methodology

An analysis of publication venues reveals a wide range of journals, underscoring the interdisciplinary character of sustainability assessment in cultural heritage. However, a few journals publish most contributions, as listed in Table 1, including those with at least 2 publications on this topic. Prominent among these are the International Journal of Life Cycle Assessment, Resources, Conservation and Recycling, Journal of Cleaner Production, and Building and Environment, all of which emphasise environmental sustainability and life-cycle perspectives. Sector-specific journals such as the Journal of Cultural Heritage, Heritage Science, and the Journal of Cultural Heritage Management and Sustainable Development also play significant roles, reflecting the importance of heritage conservation. This pattern affirms that sustainability assessment in cultural heritage is a multidisciplinary field that intersects environmental science, materials and building engineering, and heritage management.

Table 1. Distribution of papers across journals

Journal	Papers published
International Journal of Life Cycle Assessment	3
Resources, Conservation and Recycling	3
Journal of Cleaner Production	2
Journal of Cultural Heritage	2
Building and Environment	2
Cleaner Environmental Systems	2
Journal of Cultural Heritage Management and Sustainable Development	2
Journal of Building Engineering	2
Heritage Science	2
Energy Reports	2
Journal of the American Institute for Conservation	2

3.2 Content analysis

The analysis of the selected studies reveals several coherent thematic trends that demonstrate how sustainability assessment is increasingly integrated into cultural heritage conservation. First, many studies use LCA to assess the environmental implications of restoration materials and associated conservation products. Dal Pozzo et al. (2024) employ LCA to compare different consolidants used in the restoration of weathered carbonate substrates, aiming to identify formulations with reduced life-cycle impacts. Their results indicate that an innovative treatment with an aqueous diammonium hydrogen phosphate (DAP) solution outperforms commercial options. Similarly, Turk et al. (2017) assess three eco-friendly calcium carbonate-based consolidants. Their LCA study reveals that calcium acetoacetate has a lower environmental impact than the others, with 1 kg of the consolidant associated with 0.32 kg of CO2-equivalent emissions. Other articles employed LCA to assess conservation-related materials and products, including innovative nanomaterials for the preventive conservation of cultural assets (e.g. Menegaldo et al., 2023). Additional research examines the environmental effects of treatments aimed at slowing the degradation of oxygen-sensitive materials or objects in cultural heritage (e.g. Sanchez et al., 2023). Overall, these studies indicate that LCA is increasingly used as a decision-making tool for selecting restoration materials, highlighting the need to incorporate environmental performance criteria into conservation efforts alongside traditional technical and aesthetic factors.

A smaller but significant number of contributions use LCC, either alone or in combination with LCA, to analyse the economic dimension of conservation. Menegaldo et al. (2023) employ an integrated LCA–LCC methodology to assess the potential of innovative nanotechnologies for the preventive conservation of cultural objects. Their findings highlight a substantial impact linked to raw material sourcing, especially transportation. Hromada et al. (2024) apply LCC to maintenance planning in historic buildings, demonstrating that deferred maintenance increases long-term costs and risks. Their findings suggest that heritage asset management should adopt a life-cycle cost approach rather than short-term budgeting. Together, these studies emphasise the lack of economic analysis in heritage sustainability research and highlight the importance of LCC for long-term financial planning.

The social dimension remains underexplored, yet various studies broaden sustainability assessments to include social impacts. Arcese et al. (2015) aim to develop a framework for assessing social impact in the cultural heritage sector by integrating existing SLCA tools with data from social assessments of how cultural services relate to stakeholders. They propose that heritage projects can bring community benefits, including improved identity and increased employment, but also pose risks such as gentrification and unequal access. Mohaddes Khorassani et al. (2019) integrate LCA and SLCA to assess cultural heritage restoration scenarios. Their LCA findings highlight that the most impactful phases involved reconstructing materials and elements that could not be recovered because the originals were too damaged. The SLCA results show that the restoration brought

several benefits, particularly in addressing social issues within the local community. These findings indicate that incorporating social assessment into heritage sustainability analysis is still a developing area, but it remains crucial for understanding community impacts.

A final group of studies includes conceptual frameworks and systematic reviews. Flyen et al. (2020) examine how socio-cultural values can be integrated into sustainability assessments, highlighting that purely technical evaluation methods might overlook significant heritage meanings. Calabrò et al. (2024) suggest that using monitoring technologies, such as drones, can significantly improve the accuracy of initial cost estimates for cultural heritage restoration. Chelazzi et al. (2024) introduce the GREENART project, highlighting innovative, eco-friendly solutions for conserving cultural heritage, such as protective coatings, foams, and consolidants, that support broader sustainable conservation efforts. Elnaggar (2024) offers an overview of LCA applications in cultural heritage, emphasising the benefits and challenges of these life-cycle-based approaches in heritage science, and offering recommendations to guide the field towards more sustainable, eco-friendly practices.

As a result, during the content analysis phase, a systematic data extraction process was conducted. Each study was analysed to identify methodological characteristics related to the application of LCA, LCC, or SLCA, including system boundaries, functional units, data sources, and modelling assumptions, as well as the impact categories considered, the temporal horizon of the assessment, and the sustainability indicators reported. This analytical phase identified recurring sustainability indicators and decision variables across the environmental, economic, and social dimensions. These findings provided the empirical basis for developing a structured, knowledge-based, multidimensional evaluation framework to support material selection and intervention planning in cultural heritage conservation, aligned with life-cycle thinking (Table 2).

Table 2. Multi-criteria life-cycle framework

Life Cycle Assessment	Categories	Indicators	References
Environmental Life Cycle Assessment (E-LCA)	Atmospheric Pollution	Global Warming Potential (GWP)	*Dal Pozzo et al. 2024* *Elnaggar et al. 2024* *Settembre Blundo et al. 2018* *Ferrari et al. 2015* *ISO 14040-44*
		Ozone Depletion Potential (ODP)	
		Photochemical Ozone Formation Potential (POCP)	
	Resource Depletion	Water Depletion (WD)	
		Mineral Resource Depletion (MRD)	
	Ecosystems Pollution	Acidification Potential (AP)	
		Eutrophication Potential (EP)	
Life Cycle Costing (LCC)	Initial and Direct Costs	Equipment and Materials Cost (EMC)	*Settembre Blundo et al. 2018* *Menegaldo et al. 2023* *Hromada et al. 2024*
		Energy Cost (EC)	
		Human Resources Cost (HRC)	
	Post-Restoration Costs	Maintenance & Monitoring Cost (MC)	
		End-of-Life Cost (EoLC)	
Social Life Cycle Assessment (S-LCA)	Workers and Skills	Health and Safety (HS)	*Khorassani et al. 2019* *Settembre Blundo et al. 2018* *Rivera-Sánchez et al. 2014*
		Training and Education (TE)	
		Employment Stability (ES)	
		Fair Salary (FS)	
	Community and Social Benefits	Community Well-being (CW)	
		Social Inclusion and Accessibility (SIA)	
		Cultural Value (CV)	

4 Conclusions, Implications and Future Research Directions

The main contribution of this study is the definition and organisation of a multidimensional taxonomy of life-cycle–based indicators, enabling a more holistic understanding of the impacts of cultural heritage restoration interventions. It is a flexible, replicable decision-support framework adaptable to diverse cultural heritage contexts and site-specific constraints, helping to overcome the fragmented nature of existing approaches, which often focus on single dimensions of sustainability without adopting a comprehensive life-cycle perspective. The framework organises life-cycle indicators into a coherent knowledge structure that supports transparent, informed decision-making.

Building on this foundation, the framework can be further developed into an operational decision-support tool through empirical applications using multi-criteria decision-making techniques, such as the fuzzy Analytic Hierarchy Process (AHP) and the fuzzy Technique for Order Preference by Similarity to Ideal Solution (TOPSIS). These methods would allow the weighting of heterogeneous criteria and the prioritisation of alternative interventions under conditions of uncertainty and limited data availability, which are typical of cultural heritage

contexts. Through this integration, the framework can shift from a descriptive structure to a prescriptive tool that supports concrete decision-making processes.

This research design entails relevant practical and policy implications. From a managerial perspective, it would support practitioners and heritage managers in allocating resources more effectively and in identifying intervention strategies that balance performance across multiple sustainability dimensions. From a policy standpoint, it provides a robust analytical basis for developing guidelines, evaluation criteria, and funding mechanisms aligned with life-cycle sustainability principles. Moreover, by enhancing transparency and comparability, the framework can contribute to more accountable, evidence-based decision-making in the cultural heritage sector.

5 Acknowledgements

This research was supported by the PRA 2025 Call for Projects of Relevant Scientific Interest, aligned with the Strategic Research Plan of the Pegaso University.

6 Ethics Declaration

Ethical clearance was not required for this research.

7 AI Declaration

The authors did not use any AI tool to conduct this research.

References

Abbate, S., Centobelli, P., Cerchione, R., Oropallo, E., Riccio, E., 2023. Investigating Healthcare 4.0 Transition Through a Knowledge Management Perspective. IEEE Trans. Eng. Manag. 70, 3297–3310. https://doi.org/10.1109/TEM.2022.3200889

Angrisano, M., Nocca, F., Scotto di Santolo, A., 2024. Multidimensional Evaluation Framework for Assessing Cultural Heritage Adaptive Reuse Projects: The Case of the Seminary in Sant'Agata de' Goti (Italy). Urban Sci. 8, 50. https://doi.org/10.3390/urbansci8020050

Arcese, G., Di Pietro, L., Guglielmetti Mugion, R., 2015. Social Life Cycle Assessment Application: Stakeholder Implication in the Cultural Heritage Sector, in: Muthu, S.S. (Ed.), Social Life Cycle Assessment, Environmental Footprints and Eco-Design of Products and Processes. Springer Singapore, Singapore, pp. 115–146. https://doi.org/10.1007/978-981-287-296-8_4

Calabrò, F., Minniti, G.E., Fotia, A., Pucinotti, R., 2024. Recovery of historic real estate: Life Cycle Costing and economic feasibility. Procedia Struct. Integr. 64, 1759–1766. https://doi.org/10.1016/j.prostr.2024.09.181

Centobelli, P., Cerchione, R., Esposito, E., 2017. Environmental sustainability in the service industry of transportation and logistics service providers: Systematic literature review and research directions. Transp. Res. Part Transp. Environ. 53, 454–470. https://doi.org/10.1016/j.trd.2017.04.032

Cerchione, R., Esposito, E., 2016. A systematic review of supply chain knowledge management research: State of the art and research opportunities. Int. J. Prod. Econ. 182, 276–292. https://doi.org/10.1016/j.ijpe.2016.09.006

Chelazzi, D., Poggi, G., Baglioni, P., 2024. The GREENART Project: Environmentally Friendly and Low Impact Materials for the Conservation of Cultural Heritage. Stud. Conserv. 69, 25–34. https://doi.org/10.1080/00393630.2024.2342656

Chen, K., Tan, X., 2025. Does the preservation policy of historical and cultural heritage promote economic growth? Evidence from China. J. Cult. Herit. 74, 157–165. https://doi.org/10.1016/j.culher.2025.06.004

Dal Pozzo, A., Masi, G., Sassoni, E., Tugnoli, A., 2024. Life cycle assessment of stone consolidants for conservation of cultural heritage. Build. Environ. 249, 111153. https://doi.org/10.1016/j.buildenv.2023.111153

Elnaggar, A., 2024. Nine principles of green heritage science: life cycle assessment as a tool enabling green transformation. Herit. Sci. 12, 7. https://doi.org/10.1186/s40494-023-01114-z

Flyen, A.-C., Flyen, C., Fufa, S.M., 2020. Life cycle analyses applied to historic buildings: Introducing socio-cultural values in the calculus of sustainability, in: Rehabend. Presented at the REHABEND, University of Cantabria - Building Technology R&D Group, pp. 2326–2335.

Greenhalgh, T., Peacock, R., 2005. Effectiveness and efficiency of search methods in systematic reviews of complex evidence: audit of primary sources. BMJ 331, 1064–1065. https://doi.org/10.1136/bmj.38636.593461.68

Menegaldo, M., Livieri, A., Isigonis, P., Pizzol, L., Tyrolt, A., Zabeo, A., Semenzin, E., Marcomini, A., 2023. Environmental and economic sustainability in cultural heritage preventive conservation: LCA and LCC of innovative nanotechnology-based products. Clean. Environ. Syst. 9, 100124. https://doi.org/10.1016/j.cesys.2023.100124

Mohaddes Khorassani, S., Ferrari, A.M., Pini, M., Settembre Blundo, D., García Muiña, F.E., García, J.F., 2019. Environmental and social impact assessment of cultural heritage restoration and its application to the Uncastillo Fortress. Int. J. Life Cycle Assess. 24, 1297–1318. https://doi.org/10.1007/s11367-018-1493-1

Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., Mulrow, C.D., Shamseer, L., Tetzlaff, J.M., Akl, E.A., Brennan, S.E., Chou, R., Glanville, J., Grimshaw, J.M., Hróbjartsson, A., Lalu, M.M., Li, T., Loder, E.W., Mayo-Wilson, E., McDonald, S., McGuinness, L.A., Stewart, L.A., Thomas, J., Tricco, A.C., Welch, V.A., Whiting, P., Moher, D., 2021. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ n71. https://doi.org/10.1136/bmj.n71

Sanchez, S.A., Nunberg, S., Cnossen, K., Eckelman, M.J., 2023. Life cycle assessment of anoxic treatments for cultural heritage preservation. Resour. Conserv. Recycl. 190, 106825. https://doi.org/10.1016/j.resconrec.2022.106825

Turk, J., Pranjić, A.M., Tomasin, P., Škrlep, L., Antelo, J., Favaro, M., Škapin, A.S., Bernardi, A., Ranogajec, J., Chiurato, M., 2017. Environmental performance of three innovative calcium carbonate-based consolidants used in the field of built cultural heritage. Int. J. Life Cycle Assess. 22, 1329–1338. https://doi.org/10.1007/s11367-017-1260-8

Vadrucci, M., 2025. Sustainable Cultural Heritage Conservation: A Challenge and an Opportunity for the Future. Sustainability 17, 584. https://doi.org/10.3390/su17020584

Opening the Black Box of Incubation: Governance Archetypes and the Moderating Role of Regional Innovation Ecosystems in Start-up Growth

Antonio Sonetto, Silvia Tommaso, Antonio Ricciardi
University of Calabria, Arcavacata di Rende (CS), Italy

Abstract

Start-ups are widely recognised as key drivers of national competitiveness, contributing to economic growth and technological progress through their ability to innovate, adapt, and generate new markets. However, despite their strategic relevance, they face substantial barriers in the stages of development, including limited financial, organisational, and managerial resources, which constrain their ability to successfully commercialise innovations. As a result, firm survival remains a critical challenge, with a large share of start-ups exiting within a few years and only a minority achieving long-term sustainability. In this context, business incubators (BIs) have become widely adopted policy and managerial instruments aimed at supporting start-ups' development. They provide a combination of tangible and intangible resources, including physical infrastructure, advisory services, managerial support, and access to external networks, with the objective of improving start-up survival and growth prospects. However, despite their diffusion, existing research provides mixed evidence regarding their actual effectiveness, often conceptualising incubation as a *"black box"* with heterogeneous and context-dependent outcomes. This study addresses this gap by examining how public and private BI governance archetypes and regional innovation ecosystem development jointly shape start-up performance. Using a sample of 659 innovative start-ups hosted in 62 certified BIs in Italy, the analysis employs OLS regression models with clustered standard errors, integrating firm-level data from the Italian Business Register (IBR) and AIDA database for the period 2023-2024. The findings indicate that start-ups supported by private BIs achieve significantly higher revenue growth compared to those hosted in public BIs. However, this performance gap is not uniform across contexts. The level of regional innovation development positively moderates the effectiveness of public BIs, reducing their performance disadvantage in more advanced ecosystems. While private BIs exhibit robust performance regardless of territorial conditions, public BIs appear more dependent on external ecosystem quality. Finally, this research provides implications for managers, entrepreneurs, and policymakers.

Keywords: business incubators, performance, start-ups, institutional context, quantitative study

Paper type: Academic Research Paper

1 Introduction

Start-ups are key to a nation's competitive advantage, driving economic growth and technological innovation through their adaptability, risk-taking, and ability to create new industries with significant societal and economic impact (Audretsch *et al.*, 2020). However, start-ups face significant obstacles in bringing innovations to market due to limitations in size, resources, expertise, and organizational capabilities (Spender *et al.*, 2017). Therefore, start-up survival remains a major concern: research indicates that 40-80% of new ventures fail within the first five years (Colombelli *et al.*, 2016; Yang and Aldrich, 2012), and only 10-20% survive beyond ten years (Ebert *et al.*, 2019; Kalyanasundaram *et al.*, 2021), underscoring the challenge of achieving long-term sustainability (Wang and Zhou, 2022). In response to these challenges, business incubators (BIs) have emerged as important tools to support start-up creation and growth (e.g., Aernoudt, 2004), offering facility-related services, consulting and management support, and access to networks (Abduh *et al.*, 2007). BIs serve as a critical resource for facilitating start-ups' knowledge management (KM) processes (e.g., Gao *et al.*, 2021), encompassing the creation, sharing, and application of both tacit and explicit knowledge, and thereby enhancing their learning, innovation capacity, resilience, and overall organizational and financial performance. Therefore, BIs can be interpreted as intelligent KM infrastructures.

Despite the growing prevalence of BIs, empirical evidence on their effectiveness remains scarce and inconclusive (Bergman and McMullen, 2022; Cohen *et al.*, 2019; Hausberg and Korreck, 2020).

While several contributions document positive effects on firm survival, employment, sales growth, and innovation (e.g., Colombo and Delmastro, 2002; Löfsten and Lindelöf, 2002; Soetanto and Jack, 2016; Stokan *et al.*, 2015), others report no significant differences between incubated and non-incubated firms (e.g., Ferguson and Olofsson, 2004; Pena, 2004; Schwartz, 2013; Vásquez-Urriago *et al.*, 2016), or even negative effects on financial performance and growth (e.g., Dvouletý *et al.*, 2018; Ensley and Hmieleski, 2005; Schwartz, 2009). Lukeš *et al.* (2019) show that incubation negatively affects sales growth in the short term, while generating positive

effects in the long run. Additional evidence suggests that participation in venture incubation programs can increase the likelihood of firm closure and reduce the amount of funding raised (Schwartz, 2013; Yu, 2020).

These mixed findings highlight the need for further research to identify the conditions under which incubation generates meaningful outcomes (Braun and Suoranta, 2025). There is a need for studies that move beyond survival and examine start-up growth (Del Sarto *et al.*, 2020). Therefore, the lack of clear and consistent theories on how incubation operates has led scholars to describe BIs as a *"black box"* (Hackett and Dilts, 2004). The majority of the literature has primarily focused on tenant firms, paying limited attention to how the nature and intensity of BI support vary across different types of BIs (Breivik-Meyer *et al.*, 2020). Consequently, there is limited clarity regarding the relative benefits for venture development associated with different incubation program configurations. Moreover, a further promising explanation for these inconsistent findings is that incubation outcomes are contingent upon territorial and contextual factors, such as local economic conditions, institutional quality, industry clustering, and the maturity of regional innovation systems. Although several studies have emphasized the importance of incorporating these dimensions into empirical analyses (e.g., Ratinho and Henriques, 2010; Van Rijnsoever, 2020), systematic evidence remains scarce. Addressing these gaps, this study investigates how public and private BI governance archetypes influence start-up growth and how these effects vary across territorial contexts. Therefore, we defined the following research question:

RQ: To what extent do public and private incubator governance archetypes influence tenants' performance outcomes, and how does regional institutional thickness moderate these effects?

2 Hypothesis Development

Building on the framework proposed by Grimaldi and Grandi (2005), we distinguish among public BIs (business innovation centres) and private BIs (independent private incubators and corporate private incubators). Public BIs generally operate as non-profit organisations, as they are established by public institutions with the primary aim of fostering local economic development. In contrast, private BIs are profit-oriented entities, created and managed by private actors with explicit objectives of financial return. Prior research has found no statistically significant differences in start-up growth outcomes associated with the public or private ownership of BIs (Barbero *et al.*, 2012; Lukeš *et al.*, 2019). However, private BIs tend to be more performance and market driven than their publicly operated counterparts, which are more commonly guided by objectives related to social welfare and employment preservation (e.g., Grimaldi and Grandi, 2005). For private BIs, the success of their tenant firms, unlike in the case of public BIs, represents a necessary condition for their economic and entrepreneurial sustainability (Boschetti *et al.*, 2011). Therefore, we formulate the following hypothesis:

H_1: Start-ups supported by privately operated BIs exhibit higher revenue growth than those supported by publicly operated incubators.

Based on Bade (2026), the direct support provided by BIs, namely the services delivered internally to start-ups such as office space, targeted mentorship, and legal assistance, is encapsulated within the incubation program and primarily contingent on the intrinsic quality of the BI itself. Consequently, this dimension is not expected to vary with the level of development of the surrounding regional innovation ecosystem. In contrast, bridging mechanisms, i.e., the processes through which BIs connect start-ups to essential external resources such as networks, knowledge, and financial capital, are strongly dependent on the characteristics of the territorial context in which the BI operates. Given that the primary objective of public BIs is to foster regional development and safeguard employment by compensating for institutional voids through the creation of missing linkages and the bridging of knowledge and financial gaps (Grimaldi and Grandi, 2005; Harper-Anderson and Lewis, 2018; Mrkajic, 2017), we argue that such BIs are more externally embedded than their private counterparts, which tend to be more inward-oriented and rely predominantly on internal capabilities and privately cultivated networks. Accordingly, we posit that the effectiveness of public BIs relative to private ones is more strongly contingent upon the level of development of the surrounding territorial ecosystem. Therefore, we hypothesize:

H_2: The level of regional development positively moderates the relationship between public BI support and start-up growth performance.

3 Data and methods

3.1 Sample and data collection

This study is based on the analysis of data on BIs and start-ups located in Italy. The dataset was constructed by combining information retrieved from the Italian Business Register (IBR) and the AIDA database. Data

extraction was carried out in January 2026. The sample of BIs consists of the list of BIs certified by the Ministry of Enterprises and Made in Italy, which has also been considered in prior studies as the most reliable population of BIs in the Italian context (Bonfanti *et al.*, 2025). The total number of certified BIs included in the sample is 69. The start-up sample consists of innovative start-ups listed in the IBR, amounting to 11,742 firms. The legal requirements to obtain the qualification of innovative start-up or certified BI are illustrated in Figure 1.

Requirements to be a "certified" Business Incubator *Article 25, paragraph 5, of Italian Decree-Law No. 179 of October 18, 2012, converted with amendments by Law No. 221 of December 17, 2012*
• Availability of adequate facilities for hosting innovative start-ups, including spaces for testing, experimentation, or research • Presence of suitable equipment for start-up activities, such as ultra-broadband internet access, meeting rooms, and prototyping/testing machinery • Management by individuals with recognized expertise in business and innovation, supported by a permanent technical and managerial structure • Established collaborations with universities, research centres, public institutions, and innovation-oriented financial partners • Documented experience in supporting innovative start-ups
Requirements for qualification as an "innovative start-up" *Growth Decree 2.0 (Decree-Law No. 179/2012)*
• Incorporation as a limited liability company (e.g., S.r.l., S.p.A., S.a.p.a., or cooperative) • Main activity focused on development, production, or commercialization of high-tech goods or services, regardless of sector • Company age not exceeding 5 years, annual turnover below €5 million, and non-distribution of profits • Non-origin from extraordinary corporate transactions (e.g., mergers, demergers, or business transfers) • Evidence of innovative character through at least one of the following: significant R&D expenditure (at least 15% of the higher of total costs and total value of production), highly qualified personnel, or ownership/licensing of patents or registered software

Figure 1: Legal requirements for qualification as innovative start-up and certified BI (authors' elaboration)

Subsequently, for both samples, firm-level and financial data were retrieved from AIDA and merged with the information extracted from the IBR. While the number of BIs remained unchanged, the number of innovative start-ups decreased from 11,742 to 11,618 due to missing data.

As done in prior literature (Arcuri *et al.*, 2025; Lukeš *et al.*, 2019), we reconstructed the set of start-ups hosted within Italian BIs. Specifically, the list was generated by extracting the relevant information from the IBR, considered relatively up to date, as start-ups are required to update their records at least annually, and triangulating it with data collected from the official websites of the 69 BIs. In many cases, these websites provide updated lists of currently incubated start-ups as well as alumni. This process allowed us, on the one hand, to fill in missing information from the IBR and, on the other hand, to correct potential inaccuracies by excluding firms that were listed as incubated in the IBR but appeared as alumni on BI websites. This procedure enabled the construction of a reliable estimate of start-ups that were actively incubated at the beginning of the year. However, consistent with Lukeš *et al.* (2019), the absence of longitudinal information prevents identification of the duration of incubation.

The number of incubated start-ups thus amounted to 1,065. However, due to data availability constraints within AIDA, specifically, financial records being updated only through the 2024 fiscal year, the final analytical sample was restricted to 659 start-ups distributed across 62 BIs. 22.58% of the BIs are public, while 77.42% are private. Moreover, 11.29% of the BIs have a university affiliation. Regarding start-ups, 41.73% are hosted by public BIs and 58.27% by private ones. Additionally, 22.00% of the start-ups are incubated within BIs that have a university affiliation.

3.2 Variables and methodology

Table 1 reports the list of variables used in this study, identified on the basis of prior literature (e.g., Arcuri *et al.*, 2025; Lukeš *et al.*, 2019; Sansone *et al.*, 2020). To mitigate potential endogeneity issues arising from both self-selection (whereby more promising start-ups choose specific BIs) and cherry-picking (whereby high-performing BIs systematically select only the most promising candidates), we included several proxies for initial firm quality as control variables. By holding these factors constant, our model accounts for the observed heterogeneity at the time of entry.

Table 1: Description of the regression variables.

Type	Name	Definition	Source
Dependent	Tenant_growth	log (1 + tenant revenues 2024) - log (1 + tenant revenues 2023)	AIDA
Predictors	Public_Inc	Dummy variable = 1 if at least 50% of the BI's shareholders are public entities.	
	Regional_innovation_scoreboard	Mean-centered	Eurostat
Control - BIs	Inc_university_affiliation	Dummy variable = 1 if the BI is associated with a university	AIDA
	Inc_age	log (Difference between 31/12/2023 and BI founding date, divided by 365.25)	
	Inc_size	log (1 + BI total assets) as of 31/12/2023	
Control - Tenants			
	Tenant_age	log (Difference between 31/12/2023 and tenant founding date, divided by 365.25)	
	Tenant_initial_capital	log (tenant first year equity – tenant first year net income)	
	Tenant_size	log (1 + tenant total assets)	
	Tenant_qualified_team	Dummy variable = 1 if the team is composed of at least two-thirds master's degree holders, or at least one-third PhD students, PhD holders, or individuals with at least three years of certified research experience	IBR
	Tenant_R&D_intensity	Dummy variable = 1 if R&D expenditures represent at least 15% of the higher value between total production value and total costs	
	Tenant_intellectual_property	Dummy variable = 1 if the firm holds or licenses industrial property rights or owns registered software	
	Tenant_female_dominant_team	Dummy variable = 1 if the average of (i) percentage of female shareholders and (ii) percentage of female board members exceeds 50%	
	Tenant_young_dominant_team	Dummy variable = 1 if the average of (i) percentage of young shareholders and (ii) percentage of young board members exceeds 50%	
	Tenant_foreign_dominant_team	Dummy variable = 1 if the average of (i) percentage of foreign shareholders and (ii) percentage of foreign board members exceeds 50%	
	Regional_FE	BI region (NUTS 2 Italy)	AIDA
	Industry_FE	Tenant sector (NACE 2)	

Given the unavailability of panel data and the fact that the dependent variable is continuous rather than binary, the hypotheses are tested using an OLS regression model (Angrist and Pischke, 2009). Moreover, since multiple tenants are located within the same BI, the observations are not independent. Therefore, we employ BI-clustered standard errors to obtain valid and reliable p-values.

4 Preliminary results

Table 2 presents a descriptive statistical summary of the quantitative variables.

Table 2**:** Descriptive statistics for quantitative variables

Variables	Mean	SD	Min	Max
Tenant_growth	1.088652	3.631955	-11.69546	12.93785
Regional_innovation_scoreboard	-.1162569	8.018221	-21.00897	10.06103
Inc_age	2.566976	.8545125	-.0894411	4.136999
Tenant_age	.5995083	.7279528	-3.335633	2.01841
Tenant_initial_capital	9.604024	1.985973	0	14.77981
Tenant_size	11.53733	1.526032	5.451038	15.18722
Inc_size	15.5847	2.350215	11.3923	20.81247

Moreover, we conducted a correlation analysis to examine potential relationships among the independent variables (Table 3).

Table 3**:** Correlation matrix

		1	**2**	**3**	**4**	**5**	**6**	**7**
1	Public_Inc	1						
2	Regional_innovation_scoreboard	0.2192*	1					
3	Inc_university_affiliation	0.6202*	-0.0044	1				
4	Tenant_age	0.0092	-0.0908*	0.0157	1			
5	Inc_age	0.3665*	0.3741*	0.1288*	0.0361	1		
6	Tenant_initial_capital	0.0614	0.0761	0.0995*	-0.0331	0.0173	1	
7	Tenant_size	0.0930*	0.0613	0.1385*	0.3501*	0.0772*	0.4510*	1
8	Tenant_qualified_team	-0.0125	-0.0526	-0.0334	-0.0566	0.0258	0.0736	-0.0101
9	Tenant_female_dominant_team	-0.0193	-0.0031	-0.0705	-0.0498	0.0003	-0.0468	-0.1348*
10	Tenant_young_dominant_team	0.0703	0.0219	0.0390	-0.1643*	0.0438	-0.0655	-0.1082*
11	Tenant_foreign_dominant_team	0.0197	0.0935*	0.0617	-0.0178	0.0852*	-0.0323	-0.0331
12	Tenant_R&D_intensity	0.1118*	0.0558	0.1025*	0.0572	0.0504	-0.0790*	-0.0208
13	Tenant_intellectual_property	-0.0166	0.0281	-0.0052	0.0672	-0.0271	0.0233	0.0868*
14	Inc_size	0.4245*	0.3846*	-0.2056*	-0.0318	0.5905*	0.0501	0.0594
…….								
		8	9	10	11	12	13	14
8	Tenant_qualified_team	1						
9	Tenant_female_dominant_team	-0.0459	1					
10	Tenant_young_dominant_team	0.0863*	-0.0180	1				
11	Tenant_foreign_dominant_team	-0.0061	0.0115	-0.0558	1			
12	Tenant_R&D_intensity	-0.5097*	0.0635	-0.0698	0.0249	1		
13	Tenant_intellectual_property	-0.5855*	-0.0313	-0.0385	-0.0399	-0.1975*	1	
14	Inc_size	0.0414	0.2393	0.4360	-0.0131	0.0223	-0.0312	1

*$p < 0.05$

The results indicate that most correlation coefficients are low to moderate. The highest statistically significant coefficient is 0.6202 (between Public_Inc and Inc_university_affiliation), which remains below the critical threshold of 0.70; similarly, the corresponding VIF value is below the commonly accepted cutoff of 5 (4.33). This relatively higher correlation can be explained by the Italian context, where the majority of university-affiliated BIs are public in nature. Despite this correlation, the inclusion of both variables in the model is justified. From a theoretical perspective, they capture two distinct mechanisms of start-up support: while Public_Inc reflects the BI's governance structure and strategic mission, Inc_university_affiliation captures access to specific knowledge networks, R&D infrastructure, and technology transfer capabilities. From an empirical standpoint, retaining both variables allows us to isolate the *"pure"* effect of public governance on start-up growth, holding academic embeddedness constant. Excluding Inc_university_affiliation would lead to omitted variable bias, conflating the impact of public ownership with the benefits derived from academic resources. Additionally, a moderate and statistically significant negative correlation is observed between Tenant_qualified_team and Tenant_intellectual_property (−0.5855), as well as between Tenant_qualified_team and Tenant_R&D_intensity (−0.5097). This pattern can be explained by the institutional framework defining innovative start-ups in Italy, where these three variables represent alternative eligibility criteria. By law, firms are required to satisfy at least one of these conditions to be classified as innovative start-ups. This reporting behaviour generates a statistical substitution effect in the IBR data, thereby explaining the observed negative correlations. Finally, a moderate-to-low positive correlation is found between Public_Inc and Inc_age (0.3665). This relationship reflects the fact that public BIs tend to be older, as they are funded by public (national or regional) resources, which ensure greater financial stability and longevity compared to private BIs that are more exposed to market dynamics.

Table 4 reports the results of the econometric models. Overall, they show a progressive increase in explanatory power as additional controls and fixed effects are introduced.

With respect to H1, the coefficient associated with Public_Inc is consistently negative and statistically significant across most specifications. In particular, in Models (2), (4), (5), and (6), the coefficient ranges from −2.10 to −0.80, with significance levels between 1% and 10%: the magnitude of the effect decreases when additional controls are included, but remains statistically significant, confirming the robustness of the result. Therefore, H_1 is supported, suggesting that privately managed BIs are more effective in fostering start-up growth, potentially due to stronger market orientation, efficiency incentives, and performance-driven selection mechanisms.

Regarding H_2, Model (6) introduces both the Regional_innovation_scoreboard and its interaction with Public_Inc. While the direct effect of Regional_innovation_scoreboard is not statistically significant, the interaction term (Public_Inc × Regional_innovation_scoreboard) is positive and statistically significant at the 5% level (β = 0.086). This indicates that the negative effect associated with public BIs is attenuated as the level of regional innovation ecosystem development increases. Therefore, in more developed regions, publicly operated BIs become relatively more effective in supporting start-up growth. This finding provides support for H_2, highlighting a moderating effect whereby the institutional and economic environment enhances the performance of public BIs. This result can be interpreted through the distinction between internally provided services and externally oriented bridging mechanisms: while the former are largely independent of the regional context, the effectiveness of the latter, on which public BIs rely more heavily, is inherently contingent upon the quality of the surrounding innovation ecosystem. Accordingly, in more developed regions, where networks, knowledge flows, and financial resources are more accessible, public BIs are better able to activate these bridging mechanisms, thereby improving their relative impact on start-up growth. However, the estimated turning point for the Regional_innovation_scoreboard (» 12.23 based on coefficients from Model (6)) lies outside the observed range of the data (-21.01 and 10.06), implying that the moderating effect of Regional_innovation_scoreboard is not strong enough to reverse the performance gap between public and private BIs within the sample. Instead, it facilitates a progressive reduction of the growth disadvantage associated with public governance.

For start-ups incubated by private BIs, the quality of the regional ecosystem does not appear to significantly affect performance, as the coefficient for Regional_innovation_scoreboard (β2) from Model (6) is not statistically significant. This suggests that private BIs rely on sufficiently robust business models and well-established networks that enable them to generate growth for their tenants regardless of the surrounding regional development level, whether in less developed areas or in highly advanced innovation ecosystems.

Turning to the control variables, several relevant patterns emerge. Among BI characteristics, Inc_university_affiliation shows a positive and statistically significant coefficient in multiple models, suggesting that access to academic resources, knowledge networks, and technology transfer mechanisms positively influences start-up growth. Conversely, Inc age is negatively associated with start-up performance in some specifications, indicating that older BIs may be less dynamic or less aligned with current entrepreneurial needs. Inc size is consistently positive and highly significant, implying that larger BIs benefit from economies of scale and more extensive support services.

At the start-up level, Tenant_age exhibits a strong and consistently negative effect across all models, indicating that younger firms tend to grow faster, in line with typical life-cycle dynamics. The remaining start-up characteristics, including initial capital, size, team composition, R&D intensity, and intellectual property, do not display statistically significant effects, suggesting that, within this sample, BI-related factors play a more decisive role than internal firm attributes.

Table 4**:** OLS regression (with BI-clustered standard errors)

	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Public_Inc	.0049949 (.334296)	- 2.106529*** (.5259302)	.025536 (.315351)	- 1.600847*** (.5298798)	-.8053398* (.4747954)	-1.057044** (.43616)
Regional_innovation _scoreboard					.014874 (.0198658)	.0028431 (.0200885)
Public_Inc x Regional_innovation _scoreboard						.0864242** (.0361047)
BI characteristics						

Inc_University _Affiliation		2.423188** (.928241)		1.961707** (1.017345)	.9138285 (.6589951)	1.187968* (.6432003)
Inc_age		-.5399261** (.2215945)		-.6172467*** (.1961646)	-.1392863 (.2029197)	-.1311635 (.1925087)
Inc_size		.3818058*** (.1114791)		.3180605*** (.1051469)	.2019716** (.0928871)	.1911135** (.0908087)
Start-up characteristics						
Tenant_age			-1.198721*** (.2942011)	-1.068337*** (.2870778)	-1.148872*** (.2955172)	-1.129087*** (.3033418)
Tenant_initial_capital			.0012096 (.1015157)	-.0110018 (.1050075)	-.008282 (.0993773)	.0005436 (.1010086)
Tenant_size			.088628 (.1158086)	.0478185 (.1346679)	.0663641 (.1266059)	.0597069 (.1226776)
Tenant_qualified_team			-.2068075 (.8517925)	-.1688772 (.8636064)	-.1979628 (.8691399)	-.2281624 (.8666287)
Tenant_R&D_intensity			.3417339 (.8134266)	.3107993 (.8279176)	.3392476 (.8284904)	.3417692 (.8290144)
Tenant_intellectual_ property			-.1738522 (.7058374)	-.2129439 (.7317272)	-.1520829 (.713991)	-.1307418 (.7209306)
Tenant_female _dominant_team			-.0155094 (.5475452)	-.1342933 (.5927029)	-.0431676 (.5561532)	-.0279278 (.5536691)
Tenant_young _dominant_team			.4702246 (.354621)	.5105791 (.3620264)	.4664668 (.3475958)	.4481584 (.3485401)
Tenant_foreign _dominant_team			.4350566 (1.054268)	.0896384 (1.142679)	.3331551 (1.031206)	.3431374 (1.02782)
Regional_FE	NO	YES	NO	YES	NO	NO
Industry_FE	NO	NO	YES	YES	YES	YES
n	659	659	659	659	659	659
R^2	0.00	0.03	0.15	0.18	0.16	0.16

***p < 0.01; ** p < 0.05; * p < 0.10

5 Conclusions and Limitations

From a theoretical perspective, this study helps open the *"black box"* of incubation by moving beyond the simple incubated versus non-incubated dichotomy. By explicitly distinguishing between public and private governance models, the analysis shows that incubation outcomes vary systematically according to the strategic orientation of the BI. Moreover, the research responds to calls for greater contextualization in incubation research: by incorporating the moderating role of the regional innovation ecosystem, it demonstrates that the effectiveness of incubation is not context-neutral, but it is influenced by the institutional thickness of the surrounding environment.

For BI managers, the results highlight the importance of aligning strategies with the characteristics of the territorial context. Public BIs operating in less developed regions should not rely exclusively on local networks, which are often fragmented or incomplete. Instead, they should actively strengthen external bridging activities through inter-regional partnerships and structured open innovation programmes. At the same time, public BI management should focus not only on expanding networks but also on improving selection and support mechanisms for start-ups, mitigating potential bureaucratic constraints. Greater managerial autonomy and the adoption of performance-oriented evaluation systems could help reduce the gap with private BIs.

The findings also carry implications for policymakers. The lower performance of public BIs relative to private ones raises questions about the efficiency of public resource allocation. Policymakers should reconsider the key performance indicators used to evaluate public BIs, shifting towards more market-oriented performance indicators and encouraging more flexible, efficiency-driven management models. Moreover, innovation policy should be systemic rather than isolated: investment in incubation infrastructure must be complemented by broader regional innovation ecosystem-building interventions, including the improvement of technology transfer mechanisms, investment in innovation infrastructures and intermediaries, and the promotion of inter-firm collaboration through targeted incentives, such as funding programmes supporting innovations developed through the establishment of consortia.

This study is no without limitations. First, consistent with prior literature (e.g., Arcuri *et al.*, 2025), the analysis is affected by survivorship bias. The lack of historical data from the IBR implies that the available dataset likely overrepresents start-ups that survived until the data extraction date, while excluding those that failed in earlier stages. Second, due to data availability constraints, the absence of longitudinal information prevents the identification of the exact duration of the incubation period for individual start-ups, limiting the ability to capture dynamic effects over time (e.g., Lukeš *et al.*, 2019).

This work is currently at an early stage and is intended as a basis for further development. In future extensions, we plan to enrich the empirical framework by incorporating additional performance indicators, such as the amount and types of funding received and alternative measures of regional development. Moreover, we aim to strengthen the empirical strategy by implementing a broader set of robustness checks and alternative econometric specifications to validate the stability of the results across different modelling approaches.

6 Ethics Declaration

This study does not require an Ethics Committee Approval.

7 AI Declaration

The authors used AI-based tools to support a more accurate translation of the text.

References

Abduh, M., D'Souza, C., Quazi, A. and Burley, H.T. (2007), "Investigating and classifying clients' satisfaction with business incubator services", *Managing Service Quality: An International Journal*, Vol. 17 No. 1, pp. 74–91.

Aernoudt, R. (2004), "Incubators: Tool for Entrepreneurship?", *Small Business Economics*, Vol. 23 No. 2, pp. 127–135.

Angrist, J.D. and Pischke, J.-S. (2009), *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton University Press, Princeton.

Arcuri, M.C., Russo, I. and Gandolfi, G. (2025), "Productivity of innovation: the effect of innovativeness on start-up survival", *The Journal of Technology Transfer*, Vol. 50 No. 3, pp. 1111–1169.

Audretsch, D., Colombelli, A., Grilli, L., Minola, T. and Rasmussen, E. (2020), "Innovative start-ups and policy initiatives", *Research Policy*, Vol. 49 No. 10, p. 104027.

Bade, M. (2026), "Ecosystem conditions and the efficacy of entrepreneurial support organizations: Moderating effects on venture entry", *Strategic Entrepreneurship Journal*.

Barbero, J.L., Casillas, J.C., Ramos, A. and Guitar, S. (2012), "Revisiting incubation performance", *Technological Forecasting and Social Change*, Vol. 79 No. 5, pp. 888–902.

Bergman, B.J. and McMullen, J.S. (2022), "Helping Entrepreneurs Help Themselves: A Review and Relational Research Agenda on Entrepreneurial Support Organizations", *Entrepreneurship Theory and Practice*, Vol. 46 No. 3, pp. 688–728.

Bonfanti, A., Mion, G., Vigolo, V. and De Crescenzo, V. (2025), "Business incubators as a driver of sustainable entrepreneurship development: evidence from the Italian experience", *International Journal of Entrepreneurial Behavior & Research*, Vol. 31 No. 6, pp. 1430–1454.

Boschetti, C., Grandi, A. and Grimaldi, R. (2011), "Risorse, competenze e incubatori di impresa", *Sinergie Rivista Di Studi e Ricerche*, Vol. 61–62, pp. 327–349.

Braun, S. and Suoranta, M. (2025), "Incubating innovation: the role of incubators in supporting business model innovation", *Journal of Research in Marketing and Entrepreneurship*, Vol. 27 No. 2, pp. 255–276.

Breivik-Meyer, M., Arntzen-Nordqvist, M. and Alsos, G.A. (2020), "The role of incubator support in new firms accumulation of resources and capabilities", *Innovation*, Vol. 22 No. 3, pp. 228–249.

Cohen, S., Fehder, D.C., Hochberg, Y.V. and Murray, F. (2019), "The design of startup accelerators", *Research Policy*, Vol. 48 No. 7, pp. 1781–1797.

Colombelli, A., Krafft, J. and Vivarelli, M. (2016), "To be born is not enough: the key role of innovative start-ups", *Small Business Economics*, Vol. 47 No. 2, pp. 277–291.

Colombo, M.G. and Delmastro, M. (2002), "How effective are technology incubators?", *Research Policy*, Vol. 31 No. 7, pp. 1103–1122.

Del Sarto, N., Isabelle, D.A. and Di Minin, A. (2020), "The role of accelerators in firm survival: An fsQCA analysis of Italian startups", *Technovation*, Vol. 90–91, p. 102102.

Dvouletý, O., Longo, M.C., Blažková, I., Lukeš, M. and Andera, M. (2018), "Are publicly funded Czech incubators effective? The comparison of performance of supported and non-supported firms", *European Journal of Innovation Management*, Vol. 21 No. 4, pp. 543–563.

Ebert, T., Brenner, T. and Brixy, U. (2019), "New firm survival: the interdependence between regional externalities and innovativeness", *Small Business Economics*, Vol. 53 No. 1, pp. 287–309.

Ensley, M.D. and Hmieleski, K.M. (2005), "A comparative study of new venture top management team composition, dynamics and performance between university-based and independent start-ups", *Research Policy*, Vol. 34 No. 7, pp. 1091–1105.

Ferguson, R. and Olofsson, C. (2004), "Science Parks and the Development of NTBFs— Location, Survival and Growth", *The Journal of Technology Transfer*, Vol. 29 No. 1, pp. 5–17.

Gao, Q., Cui, L., Lew, Y.K., Li, Z. and Khan, Z. (2021), "Business incubators as international knowledge intermediaries: Exploring their role in the internationalization of start-ups from an emerging market", *Journal of International Management*, Vol. 27 No. 4, p. 100861.

Grimaldi, R. and Grandi, A. (2005), "Business incubators and new venture creation: an assessment of incubating models", *Technovation*, Vol. 25 No. 2, pp. 111–121.

Hackett, S.M. and Dilts, D.M. (2004), "A Systematic Review of Business Incubation Research", *The Journal of Technology Transfer*, Vol. 29 No. 1, pp. 55–82.

Harper-Anderson, E. and Lewis, D.A. (2018), "What Makes Business Incubation Work? Measuring the Influence of Incubator Quality and Regional Capacity on Incubator Outcomes", *Economic Development Quarterly*, Vol. 32 No. 1, pp. 60–77.

Hausberg, J.P. and Korreck, S. (2020), "Business incubators and accelerators: a co-citation analysis-based, systematic literature review", *The Journal of Technology Transfer*, Vol. 45 No. 1, pp. 151–176.

Kalyanasundaram, G., Ramachandrula, S. and Mungila Hillemane, B.S. (2021), "The life expectancy of tech start-ups in India: what attributes impact tech start-ups' failures?", *International Journal of Entrepreneurial Behavior & Research*, Vol. 27 No. 8, pp. 2050–2078.

Löfsten, H. and Lindelöf, P. (2002), "Science Parks and the growth of new technology-based firms—academic-industry links, innovation and markets", *Research Policy*, Vol. 31 No. 6, pp. 859–876.

Lukeš, M., Longo, M.C. and Zouhar, J. (2019), "Do business incubators really enhance entrepreneurial growth? Evidence from a large sample of innovative Italian start-ups", *Technovation*, Vol. 82–83, pp. 25–34.

Mrkajic, B. (2017), "Business incubation models and institutionally void environments", *Technovation*, Vol. 68, pp. 44–55.

Peña, I. (2004), "Business Incubation Centers and New Firm Growth in the Basque Country", *Small Business Economics*, Vol. 22 No. 3–4, pp. 223–236.

Ratinho, T. and Henriques, E. (2010), "The role of science parks and business incubators in converging countries: Evidence from Portugal", *Technovation*, Vol. 30 No. 4, pp. 278–290.

Sansone, G., Andreotti, P., Colombelli, A. and Landoni, P. (2020), "Are social incubators different from other incubators? Evidence from Italy", *Technological Forecasting and Social Change*, Vol. 158, p. 120132.

Schwartz, M. (2009), "Beyond incubation: an analysis of firm survival and exit dynamics in the post-graduation period", *The Journal of Technology Transfer*, Vol. 34 No. 4, pp. 403–421.

Schwartz, M. (2013), "A control group study of incubators' impact to promote firm survival", *The Journal of Technology Transfer*, Vol. 38 No. 3, pp. 302–331.

Soetanto, D. and Jack, S. (2016), "The impact of university-based incubation support on the innovation strategy of academic spin-offs", *Technovation*, Vol. 50–51, pp. 25–40.

Spender, J.-C., Corvello, V., Grimaldi, M. and Rippa, P. (2017), "Startups and open innovation: a review of the literature", *European Journal of Innovation Management*, Vol. 20 No. 1, pp. 4–30.

Stokan, E., Thompson, L. and Mahu, R.J. (2015), "Testing the Differential Effect of Business Incubators on Firm Growth", *Economic Development Quarterly*, Vol. 29 No. 4, pp. 317–327.

Van Rijnsoever, F.J. (2020), "Meeting, mating, and intermediating: How incubators can overcome weak network problems in entrepreneurial ecosystems", *Research Policy*, Vol. 49 No. 1, p. 103884.

Vásquez-Urriago, Á.R., Barge-Gil, A. and Modrego Rico, A. (2016), "Science and Technology Parks and cooperation for innovation: Empirical evidence from Spain", *Research Policy*, Vol. 45 No. 1, pp. 137–147.

Wang, A.X. and Zhou, K.Z. (2022), "Financial munificence, R&D intensity, and new venture survival: critical roles of CEO attributes", *Small Business Economics*, Vol. 59 No. 4, pp. 1641–1659.

Yang, T. and Aldrich, H.E. (2012), "Out of sight but not out of mind: Why failure to account for left truncation biases research on failure rates", *Journal of Business Venturing*, Vol. 27 No. 4, pp. 477–492.

Yu, S. (2020), "How Do Accelerators Impact the Performance of High-Technology Ventures?", *Management Science*, Vol. 66 No. 2, pp. 530–552.

Business Incubation as a Support Mechanism for the Circular Transition: Evidence from Italy

Antonio Ricciardi, Antonio Sonetto, Silvia Tommaso, Patrizia Pastore
University of Calabria, Arcavacata di Rende (CS), Italy

Abstract

Circular start-ups play a central role in the transition toward a Circular Economy (CE), due to their capacity to generate innovative solutions for resource recovery, waste minimisation, and regenerative value creation. However, their development is hindered by several constraints, such as limited access to financial and human resources, weak managerial skills, fragile market legitimacy, fragmented value chains, and complex regulatory environments. Within this context, business incubators (BIs) are expected to act as key enablers of entrepreneurial development, although their specific role in fostering circular entrepreneurship remains under-researched and not yet fully understood. To fill this gap, this study, based on a case analysis of a public certified Italian BI, explores how a CE-oriented incubation model supports the formation and development of circular business models, as well as the main constraints affecting its effectiveness. Primary data were collected through an interview with a BI manager and questionnaires administered to the incubated start-ups. The results show that circular incubation extends beyond traditional support mechanisms for start-ups, requiring highly customised and specialised services designed to address the specific challenges of circular start-ups. In particular, the analysed BI provides a combination of physical infrastructure, tailored business development support (especially mentoring and coaching focused on business model design and validation), financial capability enhancement, facilitation of investor relations, internationalisation support, and structured networking activities that promote industrial symbiosis and access to complementary capabilities within the ecosystem. Moreover, the findings highlight that the development of a CE-oriented incubation model has required substantial organisational and infrastructural investments, particularly to reinforce advanced mentoring competences and to upgrade physical and technical facilities suitable for experimentation and prototyping. The study advances the growing body of research on CE-focused incubation by offering empirical insights into how these models function in real-world settings. It goes beyond purely conceptual perspectives, developing a more operational understanding of how incubation processes are adapted to foster regenerative business approaches.

Keywords: business incubators, circular economy, circular start-ups, circular economy-oriented incubation model, case study

Paper type: Academic Research Paper

1 Introduction and theoretical background

Circular start-ups play a crucial role by developing innovative circular business models (e.g., Henry *et al.*, 2020; Sharma *et al.*, 2024), aimed at narrowing, slowing, closing, and regenerating resource flows (Bocken and Geradts, 2022; Konietzko *et al.*, 2020), and characterized by greater flexibility than incumbent firms (Han *et al.*, 2023; Henry *et al.*, 2020). These start-ups generate solutions that can be scaled, adopted, or integrated by established firms, enabling systemic circular transformation. However, they face significant challenges in designing scalable, replicable, and economically viable circular business models, as well as in their implementation (e.g., Gaurav *et al.*, 2019; Sharma *et al.*, 2024). Their growth is constrained by limited access to resources and capabilities (e.g., Baldassarre *et al.*, 2020) and difficulties in securing funding, as investors often remain sceptical, reinforcing financing gaps in circular transitions (Kahupi *et al.*, 2021; Kanda *et al.*, 2024). Additional barriers arise from industry-specific constraints, business model features, and idea maturity (Kanda *et al.*, 2024; Kasana *et al.*, 2024; Van Opstal and Borms, 2023). Moreover, weak market demand, driven by low environmental awareness, dependence on incumbents, and inadequate regulation, further limits their success (Engzell and Kambanou, 2024; Kanda *et al.*, 2024; Kasana *et al.*, 2024; Van Opstal and Borms, 2023; Von Kolpinski *et al.*, 2023). Consequently, their success depends on the ability to navigate these internal and external challenges (Engzell and Kambanou, 2024; Von Kolpinski *et al.*, 2023) while mobilizing resources and capabilities effectively.

Business incubators (BIs) can support circular start-ups across planning, development, and scaling phases (Millette *et al.*, 2020). Incubation is defined as *"an interactive development process where the aim is to encourage people to start their own business and to support start-up companies in the development of innovative products"* (Aernoudt, 2004, p. 127). According to Abduh *et al.* (2007), BI services include: (i) infrastructure (workspace and shared facilities), (ii) advisory and managerial support to strengthen entrepreneurial capabilities, and (iii) access to internal and external networks connecting investors, partners, and markets.

However, traditional incubation models often lack the competencies needed to support circular business models, limiting effectiveness in areas such as material flows, reverse logistics, and industrial symbiosis (Hull *et al.*, 2021; Millette *et al.*, 2020). To address this, Millette *et al.* (2020) propose a CE-oriented incubation framework aligned with CE requirements, involving multiple stakeholders (government, academia, entrepreneurs, BI management, Non-Governmental Organizations). Unlike traditional models, it is proactive, supporting entrepreneurs in identifying sustainable opportunities. A central element is Entrepreneurial Orientation (EO), which helps select entrepreneurs capable of transforming ideas into market-ready solutions, alongside Industrial Symbiosis (IS), which fosters inter-firm collaboration to optimize resources and reduce costs through exchanges. Accordingly, conventional BIs are increasingly required to integrate CE principles, such as waste reduction, resource circularity, and product life-cycle redesign (Klofsten *et al.*, 2024). Key support mechanisms include mentoring by experts in CE, sustainable supply chains, stakeholder engagement, and impact mitigation (Kanda *et al.*, 2024; Millette *et al.*, 2020), as well as tailored training programs, regulatory support, and marketing strategies to increase consumer awareness (Bank and Kanda, 2016; Klofsten *et al.*, 2024). Similarly, Surana et al. (2020) emphasize alignment with SDGs, cross-initiative coordination, performance monitoring, and multi-level capacity building. In university business incubators (UBIs), effective circular transition further requires collaboration with faculty, CE-focused training, digital repositories, and financial incentives (Kambanou *et al.*, 2025).

Nevertheless, BI literature remains underdeveloped and fragmented (Sohail *et al.*, 2023). In particular, the role of BIs in the CE context is still insufficiently explored, highlighting the need for research that informs adaptation of incubation programmes (e.g., Kambanou *et al.*, 2025; Klofsten *et al.*, 2024). Existing studies mainly assess conventional BIs' ability to support circular initiatives, often emphasizing structural and operational limitations, while dedicated analyses of CE-oriented BIs remain scarce. Similarly, research on circular start-ups is still emerging and requires further development (Suchek *et al.*, 2022).

To address these gaps, this study analyses a CE-oriented incubation model to identify key operational levers supporting circular entrepreneurship and the constraints limiting it. Accordingly, we defined the following research question:

RQ: How do CE-oriented BIs support incubated start-ups in the development of their business models?

2 Methodology

This research adopts a descriptive-exploratory single-case study design (Yin, 2018). The empirical context is Italy, a particularly suitable setting due to its advanced CE performance. Italy ranks among Europe's leaders, with a recycling rate of 72%, well above the EU average of 58% (Circular Economy Network, 2025). It also reports a circular material use rate of 20.8%, nearly double the European average of 11.8%. To ensure a high-standard context characterized by advanced incubation services, the selected BI was drawn from the list of Italian certified BIs recognized by the Ministry of Enterprises and Made in Italy and registered in the Business Register as of 26 November 2025.

Requirements to be a "certified" Business Incubator *Article 25, paragraph 5, of Italian Decree-Law No. 179 of October 18, 2012, converted with amendments by Law No. 221 of December 17, 2012*
• Availability of adequate facilities for hosting innovative start-ups, including spaces for testing, experimentation, or research • Presence of suitable equipment for start-up activities, such as ultra-broadband internet access, meeting rooms, and prototyping/testing machinery • Management by individuals with recognized expertise in business and innovation, supported by a permanent technical and managerial structure • Established collaborations with universities, research centres, public institutions, and innovation-oriented financial partners • Documented experience in supporting innovative start-ups

Figure 1: Legal requirements for qualification as certified BI (authors' elaboration)

Case selection followed a structured procedure. A systematic analysis of Italian certified BIs was conducted using triangulated data from official websites, internal documents, public reports, and press materials, with the aim of identifying organizations explicitly oriented toward the CE. The selected case is a public BI located in Northern Italy, hereafter referred to as a Green Business Incubator (GBI). Among Italian certified BIs, GBI is a distinctive benchmark as it is the only one explicitly and exclusively focused on green and CE principles since its inception.

Data collection relied on multiple sources to enhance methodological rigor and enable triangulation, thereby reducing potential bias and increasing the robustness of the findings (Bowen, 2009). Primary data were gathered through a semi-structured interview, lasting approximately one and a half hours, with a BI manager. Additionally, questionnaires are currently being administered to the incubated start-ups. To date, the survey has been completed by three incubated start-ups. Secondary data included publicly available documentation related to the BI and incubated start-ups, such as content from their official website, published reports, and relevant press articles.

3 Preliminary Results

3.1 A brief overview of GBI

GBI operates as *the "green hub"*, hosting firms in sustainability-related sectors and supporting circular start-ups. This specialization enables knowledge exchange, sector-specific networking, and shared cognitive frameworks among firms facing similar technological and regulatory constraints. Its role is embedded in a broader public policy strategy aimed not only at start-up creation but also at territorial development (e.g., Grimaldi and Grandi, 2005). GBI supports sustainability, renewable energy, circular production, resource efficiency, and environmental technologies, rather than a single niche: over-specialization could constrain entrepreneurial emergence in the regional system. Its mission is to implement a people-centred incubation model oriented to social value creation, supporting start-up growth and local development. A key feature is the co-location of early-stage start-ups with established innovative SMEs, generating mutual benefits: SMEs provide legitimacy, networks, and industrial linkages, while start-ups contribute innovation, experimentation, and flexibility.

GBI's selection process begins with evaluation of the start-up's mission, which must show a clear sustainability orientation. Projects must also align with regional development priorities, as admission decisions are embedded in territorial policy objectives. Territorial and infrastructural compatibility is also decisive. Start-ups involving heavy processing, emissions, hazardous materials, combustion processes, or large-scale logistics are excluded regardless of technological potential. Priority is given to initiatives aligned with local resource endowments; for example, regional hydropower tradition supports firms in energy management and renewable optimisation. In contrast, large-scale solar or wind projects are constrained by land availability and landscape protection rules. Finally, strong emphasis is placed on the entrepreneurial team (EO assessment) and long-term viability. Given the early-stage nature of applications, evaluation is forward-looking and interpretative, focusing on their potential. GBI's public mandate also allows greater risk tolerance than private BIs, enabling support for less mature but promising start-ups.

GBI incubation duration consists in four-year cycles, renewable if necessary. On average, start-ups remain for three to four years, while innovative SMEs may stay longer due to their ongoing ecosystem contribution.

3.2 GBI's Perspective on Challenges in Start-ups' Circular Business Models

From the perspective of GBI, and consistently with prior literature (e.g., Kanda *et al.*, 2024; Klofsten *et al.*, 2024; Lit *et al.*, 2024), circular start-ups encounter several recurring barriers, which GBI seeks to mitigate through targeted support services.

First, many start-ups experience difficulties in clearly articulating their value proposition in terms that are understandable to economic and institutional stakeholders. In several cases, early-stage circular initiatives tend to prioritise environmental or social objectives without adequately demonstrating their underlying economic feasibility. Translating environmental intent into tangible economic outcomes, such as resource recovery, cost reductions, or identifiable revenue streams, is therefore crucial. This translation enhances the credibility and readability of start-ups from a market perspective, thereby improving their ability to attract investment and easing the resource constraints that typically affect them.

A second challenge concerns market uncertainty. Highly innovative technologies rarely face pre-existing demand conditions and often require the deliberate creation of market legitimacy and acceptance.

Third, the lack of established circular supply chains represents a further constraint. The necessary inter-organisational connections are frequently absent and must therefore be actively developed by the start-ups themselves. As a result, start-ups are required not only to develop their own products or services but also to construct an ecosystem of actors who have not traditionally regarded specific waste streams as valuable economic inputs. This process entails continuous networking, brokerage, regulatory interpretation, and

negotiation with both public and private stakeholders, making business development closely intertwined with ecosystem building.

Finally, circular start-ups face significant structural complexity arising from regulatory frameworks governing supply chains. Even in cases where waste producers are willing to transfer materials at no cost, regulatory classifications and compliance obligations, such as permits for special waste, certified transport requirements, and administrative procedures, can significantly delay transactions. These constraints are often rooted in regulatory systems designed around disposal-oriented logics rather than reuse, thereby generating bureaucratic and procedural frictions despite the existence of clear economic incentives for material exchange.

3.3 GBI's service offering

3.3.1 Physical Infrastructure

GBI provides dedicated facilities through subsidised rental schemes well below market rates. In line with its public mandate, this generates immediate liquidity advantages for incubated firms, particularly early-stage start-ups with unstable revenues and limited financial buffers, by lowering fixed costs and extending their financial runway for experimentation.

Beyond cost savings, the infrastructure has a strong functional role. The BI offers an integrated package of services-utilities, security, logistics, maintenance, and mobility support-that reduces non-core operational burdens for resident firms. This decreases managerial complexity in early growth stages and limits coordination costs typically associated with independent site selection in industrial contexts, allowing start-ups to focus on product development, partner acquisition, and market testing.

In addition, the green campus environment enhances the symbolic legitimacy of hosted start-ups, strengthening their credibility with investors, public institutions, and strategic partners (e.g., Grilli and Marzano, 2023).

3.3.2 Business Development Support

A core element of GBI's incubation model is the structured provision of business development support through coaching and mentoring, aimed at transforming early-stage ideas into viable business models. This aligns with the importance of specialised mentoring in circular incubation contexts (Kanda *et al.*, 2024; Millette *et al.*, 2020), particularly for business model design and refinement (Bank and Kanda, 2016; Klofsten *et al.*, 2024).

A distinctive feature is its dual mentoring system, combining internal and external support. Each start-up is assigned an internal mentor ensuring continuity and acting as the main contact point with the BI. This is complemented by external mentors from academia, industry, and specialised sectors, selected based on start-up specific needs. While internal mentoring provides stability and oversight, external mentoring contributes specialised, domain-specific expertise, resulting in a hybrid and tailored support structure.

Support is also differentiated by start-up maturity. Many incoming start-ups, often research-based, have strong technological foundations but limited managerial capabilities for market translation. In early stages, support therefore focuses on core entrepreneurial skills, including business planning, value proposition design, customer segmentation, revenue model definition, and go-to-market strategies. This is particularly relevant for circular start-ups, where technical feasibility does not ensure commercial viability and value creation must align with value capture and stakeholder adoption conditions (e.g., Pieroni *et al.*, 2019).

As start-ups mature, support shifts from foundational training to validation-oriented assistance. Once core business elements are established, advisory activities focus on testing assumptions, refining business models, identifying market entry barriers, and strengthening commercial pathways. This is operationalised through *"Tractor"* programmes, targeting start-ups with a minimum viable product (MVP) but limited validation. Selected start-ups near market readiness receive structured guidance on validation strategies, early customer acquisition, and milestone definition, particularly for longer development cycles. Overall, business model development is treated as an iterative, evidence-based validation process rather than a static design activity.

Incubated start-ups perceive the support received has generally been adequate to address the main pre-entry competency gaps, which primarily concern commercial, managerial, financial, and regulatory aspects.

3.3.3 Financial Support

GBI acts as an intermediary between entrepreneurial projects and multiple funding sources, providing structured financial support instruments to enhance access to external capital. This is particularly relevant for CE initiatives, which often struggle to attract investment due to perceived long payback periods, uncertain demand trajectories, and complex value chains (Kahupi *et al.*, 2021; Kanda *et al.*, 2024).

Well-structured circular and sustainable start-ups are evaluated on comparable terms to conventional firms. In this context, the ability to clearly articulate business logic and scalability is crucial. Moreover, the increasing presence of ESG-oriented investors introduces actors willing to integrate sustainability into value assessment and accept longer investment horizons in exchange for measurable environmental and social impact, complementing more risk-averse non-ESG investors.

Start-ups receive support in preparing fundraising documentation. This helps founders, often focused on technological or environmental outcomes, understand how investors assess risk, milestones, scalability, and returns.

Start-ups are prepared for interactions with investors through structured activities aimed at refining pitch materials, clarifying strategic priorities, and emphasising investment-relevant business model aspects. GBI stages investor engagement only once sufficient organisational maturity is achieved, improving communication effectiveness and reducing reputational risk.

As public BI, tenants benefit from facilitated access to regional, national, and European programmes, which often co-finance prototyping and pre-market development, particularly in early-stage, pre-commercial phases.

Incubated start-ups highlight that this type of support is fundamental for their business model development. They report relying mainly on public funding and private investors, particularly venture capital.

3.3.4 Internationalization Support

Within GBI, support for internationalization is a core function. The process starts with an internal assessment to determine whether foreign market entry is strategically appropriate and operationally feasible. Before any external action, the BI conducts a diagnostic evaluation of target market attractiveness, organisational readiness, and unresolved structural weaknesses. Internationalisation is therefore not automatic but contingent on sufficient internal coherence. This screening is particularly relevant for circular start-ups, as foreign contexts differ significantly in regulatory regimes, waste management systems, energy infrastructures, procurement frameworks, and sustainability incentives (e.g., Chabowski *et al.*, 2025).

Once strategic alignment is established, support shifts to facilitation and network activation. GBI leverages its relational capital to connect start-ups with chambers of commerce, regional intermediaries, and institutional networks that help reduce entry barriers in unfamiliar markets. This brokerage function is especially critical for early-stage start-ups, which typically lack the legitimacy and reputation needed to independently develop international partnerships. Institutional ties thus partially compensate for limited reputational capital and reduce uncertainty in early foreign market entry phases (e.g., Kanda *et al.*, 2024).

3.3.5 Networking

A central component of GBI's incubation model is its structured networking strategy, which goes beyond informal interaction and is oriented toward intentional, goal-driven collaboration. The BI actively engages national and international technical and financial networks to identify partners supporting start-up development through technology transfer, open innovation, scaling, and access to funding. These networks are systematically leveraged to translate external capabilities into concrete opportunities for resident firms, mitigating resource and scaling constraints, particularly in CE contexts. In this regard, GBI fosters both external and internal innovation networks (e.g., Soetanto and Jack, 2013).

Externally, collaborations with large industrial groups enable start-ups to access technology transfer and challenge-based innovation projects. Through these channels, start-ups gain access to industrial infrastructure, specialised expertise, mentoring, and testing environments that would otherwise be inaccessible, while incumbents benefit from start-ups' innovative capacity. The BI acts as an intermediary, coordinating objectives, facilitating resource allocation, and overseeing project development. Partnerships with universities and research centres are also key, supporting the transformation of scientific knowledge and intellectual property into market-oriented solutions and enabling validation of entrepreneurial ideas. In addition, GBI's fosters links with governmental entities. As noted in Section 3.3.3, the BI also connects start-ups with public and private investor networks to facilitate access to financial resources.

Internally, collaboration among resident firms is reinforced by the coexistence of start-ups and more mature innovative SMEs. This proximity allows established firms to provide legitimacy, technical know-how, and relational capital, strengthening early-stage start-ups and improving overall ecosystem performance.

For circular start-ups in particular, these networking dynamics also support industrial symbiosis by enabling resource exchange opportunities, co-development activities, and collaborative use of materials among resident firms and industrial partners.

Incubated start-ups report a high perceived usefulness of networking services, both with external partners and with other tenants within the BI. Firms highlight GBI's strong commitment to fostering partnerships with

external actors (both established firms and start-ups), other incubated start-ups, universities and research centres, and public administrations. Start-ups also confirm GBI's active role in organising periodic networking events to facilitate connections with external partners, including both general and one-to-one formats. In addition, the BI supports participation in national and international trade fairs and sector-specific events. Co-working spaces are consistently regarded as important, as they enable both structured and spontaneous interactions among incubated start-ups.

Two out of three start-ups report having established collaborations with other tenants: one experienced no particular difficulty in forming such partnerships, while the other highlighted concerns related to the reluctance to share sensitive information, particularly regarding core technological developments.

Regarding external partnerships, two start-ups did not report significant difficulties, whereas one pointed to excessively slow decision-making processes among external partners compared to start-up requirements for timely decision-making. In this case, GBI played an important intermediary role in mitigating the issue.

Finally, one start-up reports engaging in industrial symbiosis collaborations, for which GBI played an active role in supporting their development.

3.4 Challenges in Developing a CE Incubation Model

The manager stresses that circular start-ups do not inherently require more intensive support than conventional start-ups, as the core incubation architecture remains broadly similar. However, supporting these initiatives has required GBI to address specific operational challenges and progressively develop additional technical and organizational capabilities.

A first challenge concerns infrastructural availability. When start-ups introduce highly sustainable and innovative technologies, patents or novel ideas alone are insufficient; adequate laboratories, dedicated spaces, and specialised equipment are often required to ensure safe and effective experimentation. GBI has therefore progressively developed the capacity to invest in and configure environments suitable for technically demanding innovation trajectories, particularly where experimentation involves non-standard production conditions.

A second challenge relates to business development support. When start-ups are led by highly specialized inventors, often researchers commercializing advanced patents, it becomes difficult to identify mentors capable of providing direct technological guidance, as founders may already possess rare domain-specific expertise. The BI has therefore expanded its mentor network over time to address highly specialized cases. In parallel, it has strengthened relationships with international investors able to provide external validation for highly uncertain circular projects.

Finally, institutional constraints stem from GBI's integration within a public organization. Because support must comply with formal procedures, public calls, ranking systems, and administrative verification processes, bureaucratic delays become a significant constraint (e.g., Von Bloh, 2021). As a result, decision-making is slower due to eligibility checks, appeals, and administrative procedures. This temporal mismatch is particularly critical as start-ups require rapid decisions, timely resource allocation, and managerial flexibility. Consequently, regulatory simplification is identified as a key policy priority.

4 Conclusions

From a theoretical perspective, the study contributes to the emerging literature on CE-oriented incubation by providing evidence on how such incubation models operate in practice, moving beyond conceptual definitions toward a functional understanding of how incubation processes are reconfigured to support regenerative business logics.

From a managerial perspective, the findings indicate that supporting circular start-ups requires incubation models that integrate specialised mentoring, enable collaborative linkages within the ecosystem, and foster investor readiness. More broadly, managers of traditional BIs should interpret circular incubation as an adaptive reconfiguration of existing incubation logics rather than a simple addition of sustainability-oriented services.

For policymakers, the results highlight the need to support both start-ups and BIs, strengthening incubation processes through targeted investments. They also underline the importance of developing sustainable innovation ecosystems that foster collaboration among actors, a critical condition in the CE context. Regulatory simplification remains essential, as administrative barriers continue to slow circular start-up development despite strong technological potential.

This research is still exploratory. Its main limitation is the single-case design and the fact that current evidence reflects primarily the BI perspective. In the future, we expect to obtain additional responses from recently administered questionnaires to incubated start-ups, further deepening the analysed topics through

interviews with founders and managers. Moreover, we plan to conduct further interviews with additional members of GBI's management team to gain broader insights and complementary perspectives on the topic.

5 Acknowledgements

The contents of this paper are part of a larger research, funded under the PRIN PNRR 2022 Project, which benefits from the European Union funds – Next Generation EU, entitled: CIRCOLEARE: Context Innovation through Resource Cycling Organization: Leading to Recycling Economy. Interfirm Networks for Circular Bioeconomy in the Agrifood Sector: Evidence from the Olive Oil Industry in Southern Italy - Project Code: P2022ZRPBL – CUP: H53D23008470001.

6 Ethics Declaration

This study does not require an Ethics Committee Approval.

7 AI Declaration

The authors used AI-based tools to support a more accurate translation of the text.

References

Abduh, M., D'Souza, C., Quazi, A. and Burley, H.T. (2007), "Investigating and classifying clients' satisfaction with business incubator services", *Managing Service Quality: An International Journal*, Vol. 17 No. 1, pp. 74–91.

Aernoudt, R. (2004), "Incubators: Tool for Entrepreneurship?", *Small Business Economics*, Vol. 23 No. 2, pp. 127–135.

Baldassarre, B., Konietzko, J., Brown, P., Calabretta, G., Bocken, N., Karpen, I.O. and Hultink, E.J. (2020), "Addressing the design-implementation gap of sustainable business models by prototyping: A tool for planning and executing small-scale pilots", *Journal of Cleaner Production*, Vol. 255, p. 120295.

Bank, N. and Kanda, W. (2016), "Tenant recruitment and support processes in sustainability-profiled business incubators", *Industry and Higher Education*, Vol. 30 No. 4, pp. 267–277.

Bocken, N.M.P. and Geradts, T.H.J. (2022), "Designing Your Circular Business Model", *Stanford Social Innovation Review*, Stanford Social Innovation Review, Vol. 20, p. 3439.

Bowen, G.A. (2009), "Document Analysis as a Qualitative Research Method", *Qualitative Research Journal*, Vol. 9 No. 2, pp. 27–40.

Chabowski, B.R., Gabrielsson, P., Hult, G.T.M. and Morgeson, F.V. (2025), "Sustainable international business model innovations for a globalizing circular economy: a review and synthesis, integrative framework, and opportunities for future research", *Journal of International Business Studies*, Vol. 56 No. 3, pp. 383–402.

Circular Economy Network. (2025), *Rapporto Sull'Economia Circolare in Italia*, No. 7.

Engzell, J. and Kambanou, M.L. (2024), "Incumbents versus circular start-ups in the workwear industry: Organisational and individual drivers and barriers to a circular economy", *International Small Business Journal: Researching Entrepreneurship*, Vol. 42 No. 4, pp. 551–580.

Gaurav, J.K., Rathi, V., Burnwal, K. and Chaturvedi, A. (2019), "Green startups in India: Challenges and opportunities", *The Journal of Governance-Special Issue on the Environment*, IC Centre for Governance.

Grilli, L. and Marzano, R. (2023), "Bridges over troubled water: Incubators and start-ups' alliances", *Technovation*, Vol. 121, p. 102689.

Grimaldi, R. and Grandi, A. (2005), "Business incubators and new venture creation: an assessment of incubating models", *Technovation*, Vol. 25 No. 2, pp. 111–121.

Han, D., Konietzko, J., Dijk, M. and Bocken, N. (2023), "How do circular start-ups achieve scale?", *Sustainable Production and Consumption*, Vol. 40, pp. 363–375.

Henry, M., Bauwens, T., Hekkert, M. and Kirchherr, J. (2020), "A typology of circular start-ups: An Analysis of 128 circular business models", *Journal of Cleaner Production*, Vol. 245, p. 118528.

Hull, C.E., Millette, S. and Williams, E. (2021), "Challenges and opportunities in building circular-economy incubators: Stakeholder perspectives in Trinidad and Tobago", *Journal of Cleaner Production*, Vol. 296, p. 126412.

Kahupi, I., Eiríkur Hull, C., Okorie, O. and Millette, S. (2021), "Building competitive advantage with sustainable products – A case study perspective of stakeholders", *Journal of Cleaner Production*, Vol. 289, p. 125699.

Kambanou, M.L., Hajoary, P.K. and Lindfors, A. (2025), "Supporting start-ups in the circular economy: An analysis of university-led incubators in India", *Journal of Industrial Ecology*, Vol. 29 No. 3, pp. 997–1012.

Kanda, W., Klofsten, M., Bienkowska, D., Henry, M. and Hjelm, O. (2024), "Challenges of circular new ventures: An empirical analysis of 70 cases", *Journal of Cleaner Production*, Vol. 442, p. 141103.

Kasana, S., Chavan, M., Sedera, D., Cheng, Z. and Ganzin, M. (2024), "Unlocking circular start-ups: A model of barriers", *Business Strategy and the Environment*, Vol. 33 No. 3, pp. 2546–2577.

Klofsten, M., Kanda, W., Bienkowska, D., Bocken, N., Mian, S. and Lamine, W. (2024), "Start-ups within entrepreneurial ecosystems: Transition towards a circular economy", *International Small Business Journal: Researching Entrepreneurship*, Vol. 42 No. 4, pp. 383–395.

Konietzko, J., Baldassarre, B., Brown, P., Bocken, N. and Hultink, E.J. (2020), "Circular business model experimentation: Demystifying assumptions", *Journal of Cleaner Production*, Vol. 277, p. 122596.

Lit, F.C., Huijben, J.C., Cloodt, M.M. and Paredis, E. (2024), "Business model innovation in circular start-ups: Overcoming barriers in the circular plastics economy", *International Small Business Journal: Researching Entrepreneurship*, Vol. 42 No. 4, pp. 506–550.

Millette, S., Eiríkur Hull, C. and Williams, E. (2020), "Business incubators as effective tools for driving circular economy", *Journal of Cleaner Production*, Vol. 266, p. 121999.

Pieroni, M.P.P., McAloone, T.C. and Pigosso, D.C.A. (2019), "Business model innovation for circular economy and sustainability: A review of approaches", *Journal of Cleaner Production*, Vol. 215, pp. 198–216.

Sharma, V., Manohar, S. and Paul, V. (2024), "Exploring challenges for sustainable development among circular start-ups in India", *Environment*, *Development and Sustainability*, Vol. 27 No. 7, pp. 15565–15590.

Soetanto, D.P. and Jack, S.L. (2013), "Business incubators and the networks of technology-based firms", *The Journal of Technology Transfer*, Vol. 38 No. 4, pp. 432–453.

Sohail, K., Belitski, M. and Castro Christiansen, L. (2023), "Developing business incubation process frameworks: A systematic literature review", *Journal of Business Research*, Vol. 162, p. 113902.

Suchek, N., Ferreira, J.J. and Fernandes, P.O. (2022), "A review of entrepreneurship and circular economy research: State of the art and future directions", *Business Strategy and the Environment*, Vol. 31 No. 5, pp. 2256–2283.

Van Opstal, W. and Borms, L. (2023), "Startups and circular economy strategies: Profile differences, barriers and enablers", *Journal of Cleaner Production*, Vol. 396, p. 136510.

Von Bloh, J. (2021), "The road to evidence based applicable policies for regional entrepreneurial ecosystems", *Journal of Entrepreneurship and Public Policy*, Vol. 10 No. 1, pp. 59–77.

Von Kolpinski, C., Yazan, D.M. and Fraccascia, L. (2023), "The impact of internal company dynamics on sustainable circular business development: Insights from circular startups", *Business Strategy and the Environment*, Vol. 32 No. 4, pp. 1931–1950.

Yin, R.K. (2018), *Case Study Research and Applications: Design and Methods*, SAGE Publications.

Learning to Scale: What Mentoring Really Does for Startups

Luca Simone Macca[1], Alain Daou[1,2]
[1]American University of Beirut – Mediterraneo, Paphos, Cyprus
[2]Olayan School of Business, American University of Beirut, Lebanon

Abstract

This study examines how entrepreneurial mentoring interaction processes shape founders' interpretations and enactment of scalability during early venture development. While prior research has established that mentoring positively influences entrepreneurial outcomes, such as self-efficacy, opportunity recognition, and strategic decision-making, it has largely treated mentoring as an input and focused on its effects rather than the interactional mechanisms through which influence unfolds. At the same time, research on scalability has predominantly emphasised structural and organisational dimensions, overlooking how founders come to adopt, or resist, a scalability-oriented mindset. Addressing this gap, the present study conceptualises mentoring as a sensegiving–sensemaking process through which meanings of scalability are constructed, negotiated, and translated into action. The study draws on a qualitative multiple-case design involving five early-stage startups participating in a structured pan-European mentoring program. The analysis is based on twenty mentoring sessions, complemented by follow-up interviews and secondary data. This dataset enables an in-depth examination of how mentoring unfolds over time as an interactional process. The findings show that mentoring introduces a conceptualised scalability evaluative regime, which emerged as a set of externally oriented criteria that redefine what counts as legitimate progress, viable growth, and credible venture development. This regime operates by shifting attention toward investor-oriented metrics, reducing dependence on founder time, enforcing representational clarity, and emphasising narrative coherence and distribution channels. However, findings from the founders' perspective show that founders do not merely accept these criteria passively. Instead, they engage in boundary work, actively interpreting mentoring inputs through four recurring responses: acceptance, negotiation, identity-protective resistance, and deferral. These interaction dynamics give rise to three distinct scalability trajectories: (1) relabeling, where existing activities are reframed as evidence of scalability; (2) systematising, where ventures are restructured into bounded, repeatable systems; and (3) narrativising, where growth is pursued through strengthened positioning and legitimacy-building channels. The study demonstrates that scalability is not simply a structural capability but an interpretive accomplishment emerging through ongoing interaction. By shifting attention from outcomes to processes, this research contributes to entrepreneurship literature by integrating mentoring, entrepreneurial cognition, and scalability.

Keywords: Entrepreneurial mentoring; entrepreneurial learning; scalability; early-stage startups.

Paper type: Academic Research Paper

1 Introduction

In contemporary entrepreneurial ecosystems, early-stage ventures are increasingly exposed to strong pressures to demonstrate scalability, legitimacy, and investment readiness at very early stages of development (Stein et al., 2026; Spigel, 2017; Troise et al., 2022). Mentorship programs have emerged as a central mechanism through which startups are guided in navigating these pressures, particularly in accelerator- and university-affiliated settings (Eesley and Lee, 2021; Lall et al., 2023; Sarkar et al., 2019; Seitz et al., 2024). Prior research positions mentoring as a key mechanism through which early-stage founders gain access to experiential knowledge, strategic feedback, and social capital, thereby helping them cope with uncertainty and resource constraints (Yusubova & Knoben, 2025; Wu & Tsao, 2025). Empirical studies consistently show that founders value mentoring highly and often cite it as a primary reason for participating in structured support programs (Kirsch et al., 2025; Yusubova & Knoben, 2025).

The dominant stream of mentoring research adopts an outcome-oriented perspective. Quantitative, experimental, and survey-based studies link mentoring to outcomes such as entrepreneurial self-efficacy, opportunity recognition, strategic goal attainment, and major strategic decisions including pivoting (Kirsch et al., 2025; Pereyra et al., 2021; St-Jean and Tremblay, 2020). Recent work further demonstrates that the framing of mentor feedback matters, showing that concrete feedback increases the likelihood that entrepreneurs act on advice compared to more abstract guidance (Welbourne & Miller, 2026). While these studies establish that mentoring can influence entrepreneurial behaviour, their findings are often context-dependent, and similar mentoring interventions frequently produce heterogeneous outcomes across ventures.

To address these limitations, a growing body of research conceptualises mentoring as a relational and cognitive process embedded in ongoing interaction. For example, drawing on socially situated cognition and sensemaking perspectives, Kaffka et al. (2021) highlight how feedback from mentors and other stakeholders can

disrupt entrepreneurs' existing assumptions and trigger new sensemaking through reframing, questioning, and reinterpretation. Process-oriented studies of entrepreneurial learning similarly show that entrepreneurs dynamically shift between exploratory and exploitative learning strategies in response to uncertainty and feedback, with learning trajectories shaped by situational conditions and interactional contexts (El-Awad, 2023; McKevitt and Marshall, 2015). A similar stream of research on sustainability-driven enterprises further demonstrates that entrepreneurs engage in different patterns of sensemaking depending on the ambiguity and urgency of the issues they face (Zeng et al., 2026).

Despite these advances, mentoring interactions themselves remain underexplored. Most studies rely on retrospective accounts, experiments, or surveys, rather than examining mentoring interactions as they unfold over time. This aspect follows an evident need to investigate and learn more about how the mentor-mentee relationship transforms in scalability and business model transformation. While research on scaling, pivoting, and platform growth identifies key drivers of venture growth and strategic reorientation (Granot et al., 2026; Macca et al., 2025; Nagy et al., 2025), it largely focuses on outcomes rather than on how founders learn to think in scalable terms. This lack of knowledge is especially evident for early-stage ventures, where scalability requires cognitive, strategic, and identity transformations (Liu et al., 2026; Somaya and You, 2024). In this regard, the literature points to a need for longitudinal, interaction-level research that examines how mentoring shapes founders' sensemaking and scalability thinking over time.

Building on the above literature, this study conceptualises mentoring as an interactional process through which meanings about venture development are constructed, negotiated, and enacted. Specifically, it explores how mentoring interactions influence how founders interpret and operationalise scalability during early venture development. Drawing on a qualitative multiple-case study of five startups participating in a pan-European mentoring program, this study analyses twenty mentoring sessions to uncover how mentors introduce evaluative criteria and how founders respond to them over time. Each startup was supported through multiple semi-structured mentoring sessions, generating a rich dataset of recorded and transcribed interactions that captures both mentor interventions and founder responses as they unfolded. Data collection was also supported by documentary and observational evidence (e.g., mentoring policy statements, guidelines and procedures, company websites and business plans), which were useful in validating the startup's development phase and identifying milestones.

Our preliminary findings show that: (1) mentoring consistently installed what we conceptualise as a scalability evaluative regime. Through repeated interactions, mentors introduced external evaluative criteria aligned with investor, partner, and market expectations, thereby shifting attention away from internally meaningful indicators such as effort, creativity, or early traction; (2) founders did not respond uniformly to this evaluative regime: instead, they engaged in active sensemaking and boundary work as they interpreted and translated mentoring input into action; (3) over successive interactions, these sensemaking responses stabilised into three distinct emergent scalability trajectories (i.e., relabeling trajectory, systematising trajectory, narrativising trajectory), which illustrate that scalability is not a singular outcome but an emergent process shaped by how evaluative criteria are enacted in practice.

The contribution of this study is threefold. First, it advances entrepreneurial mentoring research by providing a more nuanced understanding of why similar mentoring interventions can produce heterogeneous results across ventures. It highlights the cognitive and relational mechanisms through which scalability is constructed and offers insights into how mentoring influences strategic orientation in early-stage startups. Second, it contributes to the entrepreneurship and scalability literature by showing that scalability emerges through multiple trajectories rather than converging on a single dominant model. Third, it offers practical implications for the design of mentorship and accelerator programs by highlighting the unintended consequences of evaluative regimes and the importance of recognising founder sensemaking dynamics.

2 Theoretical Background

Entrepreneurial mentoring is widely recognised as a core component of entrepreneurial ecosystems, providing founders with access to knowledge, networks, and legitimacy (Pauwels et al., 2015). Within accelerators and incubators, mentoring functions as a key mechanism through which ventures receive strategic guidance and validation (Cavallo et al., 2025; Seitz et al., 2024). Prior research has linked mentoring to improved entrepreneurial outcomes, including enhanced self-efficacy and opportunity recognition (Baluku et al., 2019; Kirsch et al., 2025).

However, much of this literature treats mentoring as an input and evaluates its effectiveness through measurable outcomes, leaving underexplored the interactional processes through which mentoring exerts influence. Recent studies have begun to reconceptualise mentoring as a relational and cognitive process

embedded in ongoing interaction (Bergman & McMullen, 2022; Kaffka et al., 2021). In this view, mentors act as sensegivers who introduce evaluative perspectives, while founders engage in sensemaking processes that shape how these inputs are interpreted and enacted.

This perspective aligns with broader research on entrepreneurial learning, which emphasises that cognition evolves through interaction with external actors and feedback (El-Awad, 2023; Lynch & Corbett, 2023). Founders interpret mentoring advice through the lens of their prior experiences, identity, and contextual constraints, which explains variation in responses to similar mentoring interventions (Kirsch et al., 2025). Thus, mentoring effectiveness depends not only on the quality of advice but also on how it is interpreted and negotiated in interaction.

Scalability has become a central evaluative concept in entrepreneurship, particularly in technology-driven and high-growth ventures (Coviello et al., 2024). It is commonly defined as the ability to expand output or market reach without proportional increases in costs or resources (Nielsen & Lund, 2018). Achieving scalability requires alignment between technological infrastructure, organisational design, and business model logic (Granot et al., 2026; Nagy et al., 2025).

Existing research primarily examines scalability as a structural or organisational capability, focusing on business model innovation, experimentation, and ecosystem dynamics (Foss & Saebi, 2018; Sanasi, 2023). While these studies provide valuable insights into the drivers of growth, they tend to overlook the cognitive processes through which founders come to adopt a scalability-oriented mindset.

Importantly, scalability is also an evaluative construct shaped by expectations from investors, mentors, and markets (Cavallo et al., 2025). Founders must interpret these expectations and translate them into strategic decisions, often navigating tensions between short-term survival and long-term growth (Wu & Tsao, 2025). This interpretive dimension of scalability remains underexplored.

Bringing together these perspectives highlights an important gap: while mentoring research emphasises outcomes and scalability research emphasises structure, neither fully explains how founders come to interpret and enact scalability through interaction. Mentoring introduces evaluative criteria that define what constitutes a scalable venture, but founders actively negotiate these criteria rather than simply adopting them.

This study addresses this gap by conceptualising mentoring as a process through which a scalability evaluative regime is introduced and interpreted over time. This perspective enables a process-level understanding of how mentoring influences entrepreneurial cognition and how different scalability trajectories emerge from interaction dynamics.

3 Methodology

To investigate how mentoring interactions influence founders' interpretations of scalability, this study adopts a qualitative multiple-case study design. This approach is particularly suitable for examining processual phenomena and capturing the complexity of interaction dynamics over time (Eisenhardt, 1989; Yin, 2014). Unlike survey-based or experimental studies, which focus on outcomes, this design allows for in-depth analysis of how mentoring unfolds in practice. An inductive theory-building approach (Eisenhardt & Graebner, 2007; Glaser & Strauss, 1967) was employed, given the exploratory nature of our research and the limited prior theory explaining how mentoring influences interpretations of scalability (Goffin et al., 2019). By analysing multiple mentoring sessions across different ventures, the study captures both recurring patterns and meaningful variations in interaction processes.

The empirical setting consists of a structured mentoring program within a pan-European entrepreneurial support initiative. The program involved collaboration between higher education institutions, accelerators, and research organisations, providing a diverse and institutionally rich context for mentoring interactions. Five early-stage startups were selected using purposive sampling to ensure variation in business models, sectors, and developmental stages. This diversity allowed for theoretical replication and facilitated comparison across different contexts in which scalability might be interpreted differently. Each startup participated in four mentoring sessions over a three-month period, resulting in a dataset of twenty sessions. Mentors were drawn from a pool of thirteen individuals with diverse backgrounds, including entrepreneurs, investors, consultants, and academics. This diversity introduced variation in evaluative perspectives and enriched the interaction dynamics.

The primary data consist of audio and video recordings of mentoring sessions, which were transcribed and analysed. These recordings capture real-time interactions, including questions, feedback, disagreements, and moments of alignment. This level of detail enables a nuanced understanding of how mentoring unfolds as an interactional process. To complement the primary data, secondary materials were collected, including venture

documents, pitch decks, and publicly available information. Follow-up interviews were also conducted to clarify interpretations and gain additional insights into founders' perspectives.

Data were analysed inductively using grounded theory and Gioia methodology principles (Gioia et al., 2013). The analysis progressed from first-order coding to second-order themes and aggregate dimensions (Strauss & Corbin, 1990). Three key dimensions emerged: (1) the introduction of a scalability evaluative regime, (2) founder sensemaking as boundary work, and (3) emergent scalability trajectories.

4 Preliminary Results

The preliminary findings of this study indicate that entrepreneurial mentoring exerts a structured yet dynamic influence on how founders interpret and enact scalability during early venture development. Rather than functioning as a simple transfer of knowledge, mentoring interactions consistently introduce what can be conceptualised as a scalability evaluative regime, a set of externally oriented criteria that redefine how venture progress, viability, and legitimacy are assessed.

A first key insight concerns how mentoring reframes legitimacy. Across cases, mentors systematically shifted founders' attention away from internally meaningful indicators of progress, such as effort, creativity, or early traction, toward externally evaluable signals aligned with investor and market expectations. This included emphasising the distinction between activity and scalability, particularly highlighting the limitations of revenue streams dependent on founder time. Founders were encouraged to demonstrate that value creation could persist independently of their direct involvement, thereby repositioning scalability as a prerequisite for legitimacy rather than a future aspiration.

A second important finding relates to the quantification of viability. Mentoring interactions repeatedly foregrounded the importance of metrics such as customer acquisition cost, lifetime value, and measurable value propositions. Rather than treating metrics as optimisation tools, mentors framed them as gatekeeping mechanisms that determine whether a venture can be considered scalable. This introduced a shift from intuitive or narrative-driven reasoning toward evidence-based evaluation. Notably, pilots and early experiments were reframed as measurement infrastructures designed to generate quantifiable proof of value, reinforcing the expectation that scalability must be demonstrated rather than assumed.

A third insight concerns the enforcement of representational discipline. Mentors consistently challenged founders to translate broad visions into concrete, bounded representations, such as clearly defined use cases, system architectures, and investor-facing artefacts. This process reduced ambiguity and increased the legibility of ventures to external audiences. Founders were required to move from abstract descriptions of "platforms" or "solutions" to detailed explanations of how their systems function in practice. This emphasis on clarity suggests that scalability is not only a structural or operational challenge but also a representational one, dependent on how effectively ventures can be communicated and evaluated.

A fourth finding highlights the role of narrative and channels in scaling. Mentors emphasised that scalability depends on the ability to articulate a clear and compelling value proposition and to position the venture within credible distribution channels. In particular, narrative clarity was framed as a mechanism for growth, enabling ventures to be understood quickly by external stakeholders. Similarly, access to established platforms and intermediaries was presented as a means of "borrowing legitimacy," reinforcing the relational nature of scalability within entrepreneurial ecosystems.

Importantly, founders did not respond uniformly to these evaluative pressures. Instead, they engaged in ongoing sensemaking processes characterised by boundary work. Some founders accepted mentoring frames and rearticulated their ventures accordingly, adopting investor-oriented language and reframing existing activities as scalable evidence. Others negotiated these frames by adapting them to local market conditions or operational constraints, resulting in partial adoption. In certain cases, founders resisted mentoring inputs when they conflicted with core identity commitments, such as maintaining a boutique brand or preserving personalised service delivery. Finally, some founders deferred the integration of scalability criteria, prioritising immediate survival concerns and postponing evaluation to later stages.

These varied responses gave rise to distinct scalability trajectories. The relabeling trajectory involved reframing existing activities to align with scalability expectations without substantial structural change. The *systematising trajectory* entailed simplifying and formalising operations into reusable and scalable systems. The *narrativising trajectory* focused on strengthening brand coherence and leveraging channels to enhance legitimacy and growth. Together, these findings demonstrate that scalability emerges not as a uniform outcome but as a negotiated and interactionally constructed process shaped by mentoring dynamics.

5 Conclusions

This study contributes to entrepreneurship research by shifting the focus from mentoring outcomes to interaction processes. It demonstrates that mentoring influences entrepreneurial cognition through sensegiving–sensemaking dynamics rather than direct advice adoption. It also extends research on entrepreneurial cognition by highlighting the role of boundary work in interpreting external inputs (Kaffka et al., 2021).

Furthermore, the study contributes to scalability research by conceptualising scalability as an interpretive accomplishment rather than solely a structural capability. This perspective complements existing work on business model dynamics and growth (Macca et al., 2025).

For mentors, the findings emphasise the importance of making evaluative criteria explicit and recognising variation in founder interpretation. For support organisations, mentoring should be designed as an iterative process allowing reflection and adaptation. For policymakers, mentoring programs should incorporate diverse perspectives to support varied venture trajectories.

Overall, the study demonstrates that scalability is not merely a structural outcome but a socially constructed orientation emerging through interaction. By integrating mentoring, entrepreneurial cognition, and scalability, it provides a comprehensive explanation of how early-stage ventures navigate growth within entrepreneurial ecosystems.

6 Acknowledgements

This research was carried out within the framework of the *Transforming Research*, *Industry*, *& Academia into Networks for Growth & Leadership* (TRIANGLE) project (101113035). The project has received funding from the EIT Higher Education Initiative under the EIT Strategic Innovation Agenda 2021–2027. TRIANGLE is supported by EIT Urban Mobility, an initiative of the European Institute of Innovation and Technology (EIT) which is funded by the European Union. The views and opinions expressed in this manuscript are solely those of the authors and do not necessarily reflect those of the European Union or the EIT. Neither the European Union nor the EIT can be held responsible for them.

7 Ethics Declaration

No ethical clearance was required for the research conducted in this paper.

8 AI Declaration

No AI tools were used in the development or writing of this paper.

References

Baluku, M. M., Leonsio, M., Bantu, E., & Otto, K. (2018). The impact of autonomy on the relationship between mentoring and entrepreneurial intentions among youth in Germany, Kenya, and Uganda. International Journal of Entrepreneurial Behaviour & Research, 25(2), 170–192. https://doi.org/10.1108/ijebr-10-2017-0373

Bergman, B. J., & McMullen, J. S. (2022). Helping Entrepreneurs Help Themselves: A Review and Relational Research Agenda on Entrepreneurial Support Organizations. Entrepreneurship Theory and Practice, 46(3), 688–728. https://doi.org/10.1177/10422587211028736

Cavallo, A., Theodoraki, C., & Lucini-Paioni, A. (2025). Building a space entrepreneurship industry: The role of entrepreneurial support programs for startup growth. Small Business Economics. https://doi.org/10.1007/s11187-025-01135-x

Coviello, N., Autio, E., Nambisan, S., Patzelt, H., & Thomas, L. D. W. (2024). Organizational scaling, scalability, and scale-up: Definitional harmonization and a research agenda. Journal of Business Venturing, 39(5), 106419. https://doi.org/10.1016/j.jbusvent.2024.106419

Creswell, J. W., Hanson, W. E., Clark Plano, V. L., & Morales, A. (2007). Qualitative Research Designs: Selection and Implementation. The Counseling Psychologist, 35(2), 236–264. https://doi.org/10.1177/0011000006287390

Eesley, C. E., & Lee, Y. S. (2021). Do university entrepreneurship programs promote entrepreneurship? *Strategic Management Journal*, *42*(4), 833–861.

Eisenhardt, K. M. (1989). Making Fast Strategic Decisions in High-Velocity Environments. The Academy of Management Journal.

Eisenhardt, K. M., & Graebner, M. E. (2007). Theory building from cases: Opportunities and challenges. Academy of Management Journal, 50, 25–32. https://doi.org/10.5465/AMJ.2007.24160888

El-Awad, Z. (2023). Explore or exploit? Unpacking the situational conditions and cognitive mechanisms underlying entrepreneurial learning in the new venture development process. *Entrepreneurship & Regional Development*, *35*(1–2), 162–186.

Foss, N.J. and Saebi, T. (2018) Business Models and Business Model Innovation: Between Wicked and Paradigmatic Problems. Long Range Planning, 51, 9-21. https://doi.org/10.1016/j.lrp.2017.07.006

Gioia, D. A., Corley, K. G., & Hamilton, A. L. (2013). Seeking Qualitative Rigor in Inductive Research: Notes on the Gioia Methodology. Organizational Research Methods, 16(1), 15–31. https://doi.org/10.1177/1094428112452151

Glaser, B. G., & Strauss, A. L. (1967). The Discovery of Grounded Theory: Strategies for Qualitative Research. Aldine.

Granot, N., Tzafrir, S., & Tzabbar, D. (2026). Steering through the storm: Managing startup growth and pivots. *Organizational Dynamics*, *55*(1), 101205.

Guest, G., Bunce, A., & Johnson, L. (2005). How many interviews are enough? Field Methods, 18(1), 59–82. https://doi.org/10.1177/1525822x05279903

Kaffka, G. A., Singaram, R., Kraaijenbrink, J., & Groen, A. J. (2021). "Yes and. . ., but wait. . ., heck no!": A socially situated cognitive approach towards understanding how startup entrepreneurs process critical feedback. *Journal of Small Business Management*, *59*(5), 1050–1080.

Kirsch, J., Dibbern, R., Patzelt, H., & Breugst, N. (2025). Entrepreneurs' Evaluations of Mentors: The Role of Value Similarities and Mentors' Experiences. *Entrepreneurship Education and Pedagogy*, *8*(3), 371–401.

Lall, S. A., Chen, L.-W., & Mason, D. P. (2023). Digital platforms and entrepreneurial support: A field experiment in online mentoring. *Small Business Economics*, *61*(2), 631–654.

Liu, Y., Li, Y., Hao, X., & Zhang, Y. (2019). Narcissism and learning from entrepreneurial failure. *Journal of Business Venturing*, *34*(3), 496–512.

Lynch, M. P., & Corbett, A. C. (2023). Entrepreneurial mindset shift and the role of cycles of learning. Journal of Small Business Management, 61(1), 80–101. https://doi.org/10.1080/00472778.2021.1924381

Macca, L. S., Santoro, G., & Christofi, M. (2025). Managerial experimental approaches and business model dynamics: A systematic literature review. *Review of Managerial Science*.

McKevitt, D., & Marshall, D. (2015). The legitimacy of entrepreneurial mentoring. *International Journal of Entrepreneurial Behavior & Research*, *21*(2), 263–280.

Nagy, N., Bläse, R., Filser, M., Appenzeller, J., & Puumalainen, K. (2025). Platform dynamics and rapid scaling: An empirical examination of growth drivers in platform-based start-ups. *Review of Managerial Science*.

Negri, G., Ciambotti, G., Theodoraki, C., & Littlewood, D. (2025). Stagnating or flourishing? How entrepreneurial support organizations navigate constraints in nascent ecosystems. Small Business Economics, 65(3), 1651–1679. https://doi.org/10.1007/s11187-025-01042-1

Nielsen, C., Lund, M., 2018. Building scalable business models. MIT Sloan Manag. Rev. 59 (2), 65–69.

Pauwels, C., Clarysse, B., Wright, M., & Van Hove, J. (2015). Understanding a new generation incubation model: The accelerator. Technovation, 50–51, 13–24. https://doi.org/10.1016/j.technovation.2015.09.003

Pereyra, M., Aboal, D., & Rovira, F. (2021). How effective are training and mentorship programs for entrepreneurs at promoting entrepreneurial activity? An impact evaluation. *SN Business & Economics*, *1*(7), 97.

Sanasi, S. (2023). Entrepreneurial experimentation in business model dynamics: Current understanding and future opportunities. International Entrepreneurship and Management Journal, 19(2), 805–836. Scopus. https://doi.org/10.1007/s11365-023-00836-7

Sarkar, S., Osiyevskyy, O., & Hayes, L. (2019). Talking Your Way into Entrepreneurial Support: An Analysis of Satisfaction Drivers in Entrepreneur Mutual Aid Groups. Journal of Small Business Management, 57(2), 275–297. https://doi.org/10.1111/jsbm.12384

Seitz, N., Lehmann, E. E., & Haslanger, P. (2024). Corporate accelerators: Design and startup performance. Small Business Economics, 62(4), 1615–1640. https://doi.org/10.1007/s11187-023-00817-8

Somaya, D., & You, J. (2024). Scalability, venture capital availability, and unicorns: Evidence from the valuation and timing of IPOs. *Journal of Business Venturing*, *39*(1), 106345.

Spigel, B. (2017). The Relational Organization of Entrepreneurial Ecosystems. *Entrepreneurship Theory and Practice*, *41*(1), 49–72.

Stein, L., Evers, N., & O'Gorman, C. (2026). Fostering conditions for digital entrepreneurship: The role of institutional entrepreneurs in driving FinTech ecosystem development. *Technovation*, 150, 103386.

St-Jean, É., & Tremblay, M. (2020). Mentoring for entrepreneurs: A boost or a crutch? Long-term effect of mentoring on self-efficacy. *International Small Business Journal: Researching Entrepreneurship*, *38*(5), 424–448.

Strauss, A., Corbin, J.M., 1990. Basics of Qualitative Research: Grounded Theory Procedures and Techniques. Sage Publications, Inc, Thousand Oaks, CA.

Troise, C., Corvello, V., Ghobadian, A., & O'Regan, N. (2022). How can SMEs successfully navigate VUCA environment: The role of agility in the digital transformation era. *Technological Forecasting and Social Change*, *174*, 121227.

Welbourne Eleazar, M. J., & Miller, T. L. (2026). Tell me "how" and "why": Mentor feedback framing, construal level, and entrepreneurial pivoting. *Journal of Business Venturing*, *41*(1), 106549.

Wu, C.-Y., & Tsao, M.-C. (2025). By chance or by strategy? The coevolution of startups and accelerators in the case of Taiwan. *Technological Forecasting and Social Change*, *212*, 123955.

Yin, R.K., 2014. Case Study Research Design and Methods, (5th ed.). Sage, Thousand Oaks, CA. 282 pages.

Yi-En Zeng, K., Preuss, L., Thapa Karki, S., & Brieger, S. A. (2026). How Do Entrepreneurs Make Sense of and Respond to Sustainability Tensions? Insights From Sustainability-driven Enterprises. *British Journal of Management*, *37*(1), e70026.

Yusubova, A., & Knoben, J. (2025). Mentors' motives and mentoring functions: Comparing social and commercial new ventures. *European Journal of Innovation Management*, *28*(4), 1245–1268.

Accounting-Mediated Knowledge Management in Agrifood Supply Chains: Insights from Participant Observation on How Quality, Traceability and Cost Knowledge is Produced and Coordinated

Gaetano Rago
University of Salerno, Fisciano, Italy

Abstract

Agrifood supply chains are characterised by high levels of complexity, biological variability and increasing regulatory and societal pressures for transparency, traceability and sustainability. In such contexts, knowledge management (KM) plays a critical role in enabling coordination among heterogeneous actors, including farmers, processors, quality managers, accountants and auditors. However, existing literature has largely conceptualised KM in supply chains in abstract terms, focusing on information sharing and integration, while paying limited attention to the everyday practices through which knowledge is produced, translated and made actionable.

This study addresses this gap by examining how knowledge is managed in practice within an agrifood supply chain and by exploring the role of accounting and auditing artefacts in mediating this process. Drawing on an ethnographic research design based on one year of participant observation within a large Italian agrifood processing firm and its network of partners, the study provides a fine-grained analysis of the micro-practices through which knowledge is transformed into organizationally relevant forms.

The findings show that knowledge management is not a linear process of data transfer, but a situated and ongoing accomplishment. Three key mechanisms are identified. First, translational practices convert tacit and experiential knowledge into standardized categories and inscriptions, such as quality classifications, traceability records and cost analyses. Second, accounting and auditing artefacts operate as boundary objects that enable coordination across different communities of practice while shaping visibility and accountability. Third, governance routines, including audits, meetings and performance reviews, serve as arenas in which knowledge claims are validated and stabilised.

The study contributes to knowledge management literature by demonstrating that knowledge conversion is materially mediated by calculative and documentary practices. It also extends accounting research by showing that accounting functions as a knowledge infrastructure linking quality, traceability, cost and sustainability. Overall, the paper advances an accounting-mediated view of KM, highlighting the central role of accounting and auditing artefacts in making knowledge visible, credible and actionable in complex agrifood supply chains.

Keywords: Knowledge management; agrifood supply chains; accounting artifacts; Management accounting; ethnography.

Paper type: Academic Research Paper

1 Introduction

Agrifood supply chains represent one of the most complex and highly regulated organizational fields in contemporary economies. Their complexity stems from the interplay between biological processes, environmental variability, and socio-economic constraints, combined with increasing demands for food safety, traceability, and sustainability (Trienekens *et al.*, 2012; Busch, 2011). Unlike industrial supply chains, which are typically characterized by standardization and predictability, agrifood systems must continuously adapt to uncertain production conditions, product perishability, and fluctuating demand patterns. In such a context, Knowledge Management (KM) becomes central to understanding how dispersed competencies are generated, translated, and mobilized across the supply chain (Corso et al., 2010).

A significant portion of the knowledge circulating within agrifood systems is tacit and deeply embedded in practice. Assessments of field conditions, agronomic decisions, product freshness, and production constraints largely rely on professional judgment and locally situated experience (Wenger, 1999). However, such knowledge cannot easily traverse organizational boundaries unless it is translated into forms that are mobilizable within decision-making processes. Existing literature on supply chain management has extensively emphasized the importance of information sharing and integration as drivers of performance (Mentzer *et al.*, 2001; Hult *et al*, 2007). Similarly, studies on inter-organizational relationships have highlighted the role of trust, collaboration, and governance mechanisms in facilitating coordination (Fawcett *et al*., 2008; Caglio and Ditillo, 2008). However, these perspectives often conceptualise knowledge in relatively abstract terms, focusing on systems,

technologies, or formal agreements, while overlooking the everyday practices through which knowledge is produced, negotiated, and made actionable.

This gap is particularly evident in accounting research. While management accounting has increasingly been recognized as a key mechanism for coordinating inter-firm relationships (Mouritsen and Thrane, 2006; Caglio and Ditillo, 2008), limited attention has been paid to how accounting artefacts mediate knowledge flows in complex supply chain settings. Yet, in agrifood contexts, many of the objects through which coordination is achieved, such as traceability records, quality indicators, and cost reports, are inherently accounting-based or accounting-mediated. These artefacts do not simply represent reality; they actively construct it by defining categories, establishing comparability, and allocating responsibility (Miller, 2020).

The increasing emphasis on sustainability and ESG disclosure further intensifies these dynamics.

Agrifood firms are now required to demonstrate not only economic performance but also environmental and social impacts, thereby expanding the scope of knowledge that must be produced, validated, and communicated (Gray, 2010; Burritt and Schaltegger, 2010). This raises critical questions about how different forms of knowledge are integrated and governed across organizational boundaries.

Building on these considerations, the research question guiding this study is as follows: *how is knowledge managed through everyday practices within an agrifood supply chain, and what role do accounting and auditing artifacts play in translating and legitimising knowledge among heterogeneous actors?*

To answer this question, the study adopts an ethnographic approach based on one year of participant observation within a large processing firm and its supply network. By focusing on micro-level practices, the research seeks to uncover the "invisible work" through which knowledge is transformed into actionable organizational facts, shedding light on the situated processes through which coordination is practically accomplished.

The remainder of the paper is organised as follows. The next section develops the theoretical framework by integrating KM perspectives on tacit and explicit knowledge conversion (Nonaka, 1994) and communities of practice (Wenger, 1999). The methodology section outlines the ethnographic research design and analytical approach. The results and discussion section examines the mechanisms through which accounting and auditing artifacts mediate KM practices and elaborates on their implications. Finally, the concluding section discusses the limitations of the study and suggests avenues for future research.

2 Theoretical Framework

KM has been widely recognised as a critical driver of organizational effectiveness, particularly in contexts characterised by uncertainty, interdependence, and distributed expertise (Grant, 1996; Nonaka, 1994). In supply chain environments, knowledge is inherently fragmented across multiple actors, each possessing distinct capabilities, interests, and interpretive frameworks. This fragmentation is especially pronounced in agrifood supply chains, where biological variability, environmental uncertainty, and regulatory pressures require continuous coordination among heterogeneous stakeholders (Trienekens *et al.*, 2012). In such contexts, KM cannot be reduced to the mere exchange of information; rather, it involves the ongoing production, translation, and validation of knowledge that is both locally situated and organizationally actionable.

Nonaka's (1994) theory of knowledge creation provides a foundational lens for understanding these dynamics. The SECI model conceptualises knowledge as emerging from the interaction between tacit and explicit dimensions, emphasising processes of socialisation, externalisation, combination, and internalisation. In agrifood supply chains, this interplay is particularly salient, as much of the relevant knowledge, such as farmers' experiential understanding of cultivation conditions or quality inspectors' sensory assessments, remains tacit and embedded in practice. For such knowledge to travel across organizational boundaries, it must be externalised and transformed into standardized representations that can be communicated, compared, and acted upon. This process is not merely technical but inherently social and interpretive, involving negotiation and sensemaking among actors with different perspectives (Boland and Tenkasi, 1995).

Wenger's (1999) concept of communities of practice further enriches this perspective by highlighting the situated and collective nature of knowing. Knowledge is not simply possessed by individuals but enacted through participation in shared practices. Agrifood supply chains can thus be understood as constellations of overlapping communities of practice (farmers, agronomists, quality managers, procurement specialists, accountants), each with its own norms, languages, and evaluative criteria. Coordination across these communities requires mechanisms that enable alignment without erasing differences, allowing actors to collaborate despite divergent forms of expertise. Prior research has shown that such inter-organizational knowledge sharing can generate significant performance benefits (Hult *et al.*, 2007; Dyer and Nobeoka, 2000), but it is also fraught with challenges related to trust, power asymmetries, and cognitive distance (Fawcett *et al.*, 2008).

Within this complex landscape, accounting emerges as a central yet often under-theorised mechanism for mediating knowledge. Moving beyond its traditional representational view, accounting has increasingly been conceptualised as a performative practice that actively shapes organizational realities (Hopwood, 2013).

Calculative devices, such as cost reports, performance indicators, traceability records, and sustainability metrics, do not merely capture pre-existing phenomena; they construct them by defining categories, establishing comparability, and structuring attention. In doing so, accounting contributes to determining what is considered relevant, credible, and actionable knowledge within and across organizations.

This performative role is particularly evident in inter-organizational settings, where shared understandings cannot be taken for granted. Research on management control in inter-firm relationships has highlighted how accounting systems facilitate coordination by providing common metrics and evaluation frameworks (Caglio and Ditillo, 2008). Mouritsen and Thrane (2006) argue that accounting acts as a mediating instrument that enables collaboration while simultaneously embedding governance structures and power relations. In supply chains, accounting thus operates as both a coordinating and a disciplining mechanism, enabling alignment while shaping the distribution of responsibilities and decision rights.

The concept of boundary objects (Star and Griesemer, 1989) offers a particularly useful lens for understanding how accounting artefacts perform this mediating role. Boundary objects are artefacts that are sufficiently robust to maintain a common identity across different social worlds, yet flexible enough to accommodate diverse interpretations. In agrifood supply chains, artefacts such as traceability documents, quality standards, cost analyses, and audit reports function as boundary objects that enable coordination among heterogeneous actors. These artefacts allow different communities of practice to interact and align their activities without requiring full consensus, thereby supporting cooperation in contexts characterised by epistemic diversity. At the same time, they embed specific assumptions and priorities, influencing how problems are framed and how solutions are evaluated.

The increasing emphasis on traceability and sustainability further expands the scope of these dynamics. Traceability systems, initially developed to ensure food safety and regulatory compliance, have evolved into broader infrastructures for transparency and accountability (Trienekens *et al.*, 2012). They generate detailed data on product origins, production processes, and logistical flows, enabling organizations to monitor and communicate information across the supply chain. Simultaneously, the rise of sustainability accounting has extended the boundaries of what is measured and reported, incorporating environmental and social dimensions alongside economic performance (Burritt and Schaltegger, 2010; Gray, 2010). In agrifood contexts, this includes issues such as resource use, emissions, labour conditions, and ethical sourcing.

The integration of these diverse knowledge domains introduces significant challenges, as different stakeholders may hold divergent expectations regarding what constitutes relevant and legitimate knowledge. Regulators, certification bodies, consumers, and supply chain partners often demand different forms of evidence, leading to potential tensions and trade-offs. As a result, organizations must develop mechanisms to translate, align, and legitimise multiple forms of knowledge, ensuring that they are simultaneously credible, comparable, and operationally useful. Accounting and auditing artefacts play a crucial role in this process, as they provide the infrastructures through which such knowledge is stabilised and rendered auditable. Building on these perspectives, this study conceptualises agrifood supply chains as “knowledge ecologies” in which multiple forms of knowledge interact through a combination of social practices and calculative infrastructures. Within this ecology, accounting is not merely a supporting tool but a central mediating mechanism that enables the translation, circulation, and validation of knowledge across organizational boundaries. By focusing on the everyday practices through which these processes unfold, the study advances an accounting-mediated view of KM, highlighting the constitutive role of calculative artefacts in shaping coordination, accountability, and decision-making in complex supply chain environments.

3 Methodology

The study adopts a qualitative research design based on participant observation as an ethnographic field research approach (Spradley, 1980). This approach is particularly suitable for investigating everyday practices, informal coordination and the situated use of accounting and auditing artefacts in organizational life (Ahrens and Chapman, 2006). The empirical investigation was conducted from September 2024 to September 2025 within a large Italian agrifood processing firm and involved the observation of its interactions with selected upstream and downstream partners. The researcher was immersed in key organizational settings, including procurement planning, quality control, production scheduling, sustainability practices, audit preparation and performance review meetings. This allowed the study to capture both routine activities and critical episodes, such as quality deviations, traceability issues and supplier negotiations.

The researcher adopted the role of participant observer, combining direct involvement in the field with continuous analytical distance. This position enabled close access to everyday interactions, informal exchanges and documentary practices that would have been difficult to capture through interviews alone. At the same time, the researcher maintained a reflexive stance, recording not only observed practices but also contextual interpretations and potential effects of their own presence in the field. Informed consent was requested from the organization and from the actors involved in the research process. Confidentiality and anonymity were ensured by removing identifying details concerning individuals, the firm and its supply chain partners.

Data collection relied on three main sources: participant observation, informal conversations and documentary analysis. Fieldnotes were written throughout the research process to record observed interactions, contextual details and emerging interpretations (Emerson *et al.*, 2011). Informal conversations with managers, quality technicians, procurement officers, controllers and other organizational actors helped clarify the meaning of observed practices and reconstruct the rationale behind everyday decisions (Musante and DeWalt, 2010). In addition, the analysis considered relevant artefacts, including quality reports, traceability records, audit checklists, dashboards, cost and margin analyses and sustainability-related documentation.

Data analysis followed an iterative and abductive logic, moving between empirical material and theoretical concepts. Particular attention was given to critical episodes in which knowledge was contested, negotiated or stabilised. The analysis was informed by concepts such as knowledge conversion (Nonaka, 1994), communities of practice (Wenger, 1998), boundary objects (Star and Griesemer, 1989) and the performative role of accounting, while remaining grounded in the empirical material.

The methodological design also presents limitations. First, the study is based on a single organizational setting, which limits statistical generalisation. Second, participant observation may be influenced by the researcher's position, interpretations and access to specific situations. Third, some informal exchanges and sensitive organizational dynamics may have remained only partially visible. Nevertheless, prolonged field engagement, triangulation across multiple data sources and reflexive interpretation strengthened the credibility of the analysis and supported a rich understanding of how accounting and auditing artefacts mediate knowledge management in practice.

To enhance the methodological rigour, the study combined prolonged field engagement, triangulation across multiple data sources and reflexive interpretation. Although the ethnographic design limits statistical generalisation, it provides a rich understanding of the micro-practices through which accounting and auditing artefacts mediate KM in agrifood supply chains.

4 Results and Discussion

The ethnographic evidence shows that knowledge management within the observed agrifood supply chain is not enacted through the linear transfer of information, but through situated practices in which heterogeneous forms of knowledge are progressively translated, stabilised and made actionable. In line with the theoretical framework, the supply chain operates as a "knowledge ecology" in which multiple communities of practice interact and where accounting and auditing artefacts play a central mediating role. Knowledge moves from tacit, local and experiential forms to explicit, standardised and auditable representations; however, this movement is neither neutral nor automatic, as it is shaped by calculative practices, documentary requirements and governance routines.

A first finding concerns the translation of tacit knowledge into organisational facts. Observations in the receiving area and during quality inspections show that product evaluation initially relies on embodied and sensory judgement. Quality technicians assess colour, consistency, defects and freshness through experience-based evaluation before any formal classification is entered into the system. As one technician explained during an informal conversation, "you understand certain things before writing them down, but if they are not written down, they do not exist for the rest of the company". This statement captures a central feature of the observed setting: tacit knowledge becomes organisationally relevant only when it is translated into a recognised category, such as a quality grade, a deviation report or a traceability entry.

For example, a batch described informally as "*not uniform*" during initial inspection became actionable only once it was codified as a quality deviation within the inspection sheet. Once documented, the assessment could travel beyond the receiving area and be used by procurement to discuss supplier performance, by production to adjust processing decisions and by accounting to evaluate yield and cost implications. This pattern resonates with Nonaka's (1994) notion of knowledge conversion, but the findings refine it by showing that externalisation is not merely cognitive or communicative. It is materially mediated by artefacts that impose specific formats, categories and thresholds on knowledge. Accounting and quality-related documents therefore act as translation

devices: they make tacit knowledge mobile, but they also reshape it according to managerial, regulatory and economic logics.

A second finding concerns the role of accounting and auditing artefacts as boundary objects. Traceability records, quality reports, cost analyses and audit checklists were repeatedly mobilised by actors belonging to different communities of practice. Their value did not lie in producing a single interpretation, but in allowing different actors to coordinate around a shared object. During one traceability review, the same file was used by the quality manager to verify compliance, by the production planner to reconstruct batch sequences, by procurement to check supplier codes and by administrative staff to assess whether the documentation would be sufficient in case of a customer claim. As one manager noted informally, "*the same document speaks different languages depending on who is reading it*".

This finding supports the concept of boundary objects developed by Star and Griesemer (1989). The traceability file remained stable enough to coordinate action while flexible enough to accommodate different professional concerns. However, the fieldwork also shows that such artefacts are not neutral. They privilege what can be documented, measured and audited. What is recorded becomes defensible; what remains experiential becomes more fragile. This dynamic was evident in several informal conversations, where actors repeatedly emphasised that "*knowing is not enough; you must be able to prove it*". Accounting and auditing artefacts therefore enable coordination, but they also structure visibility, legitimacy and accountability.

A third finding relates to the performative role of accounting in reframing organisational issues. During performance review meetings, technical and operational problems were frequently translated into economic terms through cost and margin analyses. Variability in raw material quality, initially discussed as an agronomic or production issue, became a matter of yield loss, increased unit cost and margin pressure once represented in financial reports. In one observed discussion, the production manager explained the slowdown of the line by referring to the variability of incoming raw material. When the controller presented the corresponding cost impact, the discussion shifted. As one participant commented after the meeting, "*when the number appears, the problem changes level*".

This episode illustrates the performative role of accounting. The margin analysis did not merely describe an operational issue; it altered its organisational relevance. Once translated into cost and margin effects, the issue became more visible to senior management and more likely to generate corrective action. This supports the view of accounting as a calculative practice that shapes organisational attention, relevance and accountability (Hopwood, 1987). At the same time, the findings show the ambivalence of this process: accounting makes problems governable, but it may also narrow attention to what can be economically quantified.

A fourth mechanism concerns governance routines. Meetings, audits, performance reviews and informal cross-functional interactions were the arenas in which knowledge claims were tested, negotiated and stabilised. Documents alone did not produce certainty. They had to be interpreted, questioned and connected to situated explanations. During audit preparation, for instance, the quality team did not simply collect evidence; it reconstructed sequences of events, checked dates, compared forms and added explanations to make the documentation intelligible to external evaluators. As one quality manager observed, "*the auditor does not only want the document; they want to understand the story behind the document*".

This finding is particularly important because it shows that auditability is not equivalent to data availability. Auditability requires coherence, chronology and credibility. The organisation must be able to demonstrate not only that procedures were followed, but also that events can be reconstructed in a plausible and verifiable way. Governance routines therefore function as sites of knowledge validation, where documentary evidence and situated explanations are brought together. This extends Wenger's (1998) view of knowing as socially embedded by showing that, in regulated supply chains, knowledge becomes legitimate through collective practices of verification and justification.

Finally, the analysis shows that sustainability knowledge entered the supply chain through the same accounting-mediated mechanisms. Sustainability was not observed as a separate domain detached from operational and accounting practices. Rather, it appeared in supplier evaluation, quality standards, certifications, customer requirements and cost discussions. In one informal conversation, a procurement manager stated that "*price is no longer the only variable; now every supplier also brings a documentation burden*". Similarly, a controller noted that "*sustainability requirements always arrive with a cost implication, even when they are presented as values*".

These observations suggest that sustainability, quality, traceability and cost are not separate knowledge domains in practice. They are bundled together and made governable through the same calculative and documentary infrastructures. A supplier decision may simultaneously concern origin, certification, price, reliability and environmental or social requirements. This confirms that sustainability accounting expands the

scope of organisational knowledge beyond traditional financial performance (Gray, 2010; Burritt and Schaltegger, 2010), while also showing that such expansion must be integrated into existing routines of calculation, documentation and accountability.

Overall, the findings suggest that transparency in agrifood supply chains is not produced simply by collecting more information. It is a practical accomplishment generated through translation, documentation, discussion and validation. Traceability systems and accounting reports provide the infrastructure, but transparency emerges only when actors can connect inscriptions to events, explanations and responsibilities. In this sense, knowledge management in agrifood supply chains is fundamentally accounting mediated. Accounting and auditing artefacts stabilise knowledge, make it transferable across communities of practice, and define the conditions under which it becomes visible, credible and actionable.

Theoretically, the study contributes to KM literature by showing that knowledge conversion in supply chains is materially mediated. While Nonaka's (1994) framework explains the movement between tacit and explicit knowledge, and Wenger's (1998) communities of practice illuminate the social embeddedness of knowing, this study shows that knowledge becomes actionable through accounting and auditing artefacts that format and legitimise it. The study also contributes to accounting research by extending the role of accounting in inter-organizational relationships. Prior studies show that accounting supports coordination and control across firm boundaries (Mouritsen *et al*., 2001; Caglio and Ditillo, 2008). This study adds that accounting also operates as a knowledge infrastructure, linking operational events, traceability, sustainability and financial consequences.

Practically, for managers the findings indicate that improving knowledge management in agrifood supply chains does not simply mean introducing more sophisticated digital tools or increasing the quantity of data collected. It requires designing routines and artefacts that support meaningful translation across professional communities. Quality reports, cost analyses, traceability records and sustainability indicators should be treated as interconnected devices rather than isolated reporting instruments. Their effectiveness depends on the capacity of actors to interpret them jointly and to connect them to situated operational knowledge. Finally, for policy makers and certification bodies, the findings suggest caution against overly formalistic approaches to transparency and accountability. Standards and reporting requirements are essential, but they must be accompanied by an understanding of the practical work through which evidence is produced. The credibility of agrifood disclosure depends not only on the presence of records, but on the organizational routines that make those records reliable, interpretable and auditable.

5 Conclusion

This study examined how knowledge is managed within agrifood supply chains and how accounting and auditing artefacts contribute to translating and legitimising knowledge among heterogeneous actors. Drawing on participant observation, the analysis shows that KM is not merely a matter of information sharing, but a situated process through which tacit, experiential and local knowledge is transformed into actionable organizational knowledge.

The study also presents some limitations that should be acknowledged. First, its ethnographic design is based on participant observation within a single large processing firm and a selected number of supply chain partners. This choice allows for an in-depth understanding of everyday practices but limits the possibility of statistical generalisation. Second, the analysis focuses primarily on accounting and auditing artefacts, while other infrastructures, such as digital platforms, informal relationships and data-driven technologies, are considered only marginally. Third, the findings are situated within a specific agrifood context and may vary across different products, institutional settings and supply chain configurations.

Future research could address these limitations by developing comparative studies across different agrifood supply chains, sectors and countries. Longitudinal research would also be valuable to examine how accounting-mediated knowledge practices evolve over time, especially in response to regulatory changes, sustainability pressures and technological innovation. Further studies could investigate how blockchain, artificial intelligence and advanced data analytics reshape traceability, sustainability accounting and inter-organizational accountability. This would help clarify whether digital technologies reinforce existing accounting-mediated knowledge practices or generate new forms of visibility, control and coordination.

Overall, this study suggests that KM in agrifood supply chains is fundamentally accounting mediated. Accounting and auditing artefacts do not simply document knowledge after it has been produced; they actively participate in making knowledge visible, credible and governable across organizational boundaries.

6 Ethics Declaration

Informed consent was obtained from the organization and the individuals involved in the study.

Confidentiality and anonymity were ensured by removing identifying details related to participants, the firm and its supply chain partners. Data were handled securely and used only for research purposes. Given the participant observation design, the researcher maintained a reflexive stance to minimise interference with everyday activities and to avoid disclosing sensitive commercial information.

7 AI Declaration

AI tools, specifically ChatGPT, were used solely to support language editing, enhance clarity, and refine the academic style of the manuscript.

References

Ahrens, T. and Chapman, C.S. (2006) "Doing qualitative field research in management accounting: positioning data to contribute to theory", *Accounting, Organizations and Society*, Vol. 31, No. 8, pp 819-841.

Boland, R.J. Jr and Tenkasi, R.V. (1995) "Perspective making and perspective taking in communities of knowing", *Organization Science*, Vol. 6, No. 4, pp 350-372.

Burritt, R.L. and Schaltegger, S. (2010) "Sustainability accounting and reporting: fad or trend?", *Accounting, Auditing & Accountability Journal*, Vol. 23, No. 7, pp 829-846.

Busch, L. (2013) *Standards: Recipes for Reality*, MIT Press, Cambridge, MA.

Caglio, A. and Ditillo, A. (2008) "A review and discussion of management control in inter-firm relationships: achievements and future directions", *Accounting, Organizations and Society*, Vol. 33, No. 7-8, pp 865-898.

Corso, M., Dogan, S.F., Mogre, R. and Perego, A. (2010) "The role of knowledge management in supply chains: evidence from the Italian food industry", *International Journal of Networking and Virtual Organisations*, Vol. 7, No. 2-3, pp 163-183.

Dyer, J.H. and Nobeoka, K. (2000) "Creating and managing a high-performance knowledge-sharing network: the Toyota case", *Strategic Management Journal*, Vol. 21, No. 3, pp 345-367.

Emerson, R.M., Fretz, R.I. and Shaw, L.L. (2011) *Writing Ethnographic Fieldnotes*, University of Chicago Press, Chicago.

Fawcett, S.E., Magnan, G.M. and McCarter, M.W. (2008) "Benefits, barriers, and bridges to effective supply chain management", *Supply Chain Management: An International Journal*, Vol. 13, No. 1, pp 35-48.

Grant, R.M. (1996) "Toward a knowledge-based theory of the firm", *Strategic Management Journal*, Vol. 17, No. S2, pp 109-122.

Gray, R. (2010) "Is accounting for sustainability actually accounting for sustainability... and how would we know? An exploration of narratives of organisations and the planet", *Accounting, Organizations and Society*, Vol. 35, No. 1, pp 47-62.

Hopwood, A.G. (2013) "The archaeology of accounting systems", in *Accounting from the Outside*, Routledge, London, pp 443-470.

Hult, G.T.M., Ketchen, D.J. and Arrfelt, M. (2007) "Strategic supply chain management: improving performance through a culture of competitiveness and knowledge development", *Strategic Management Journal*, Vol. 28, No. 10, pp 1035-1052.

Mentzer, J.T., DeWitt, W., Keebler, J.S., Min, S., Nix, N.W., Smith, C.D. and Zacharia, Z.G. (2001) "Defining supply chain management", *Journal of Business Logistics*, Vol. 22, No. 2, pp 1-25.

Miller, P. (2020) "Accounting, organizing and economizing", *Journal of Information and Management*, Vol. 40, No. 1-2, pp 9-20.

Mouritsen, J. and Thrane, S. (2006) "Accounting, network complementarities and the development of inter-organisational relations", *Accounting, Organizations and Society*, Vol. 31, No. 3, pp 241-275.

Musante, K. and DeWalt, B.R. (2010) *Participant Observation: A Guide for Fieldworkers*, Bloomsbury Publishing, London.

Nonaka, I. (1994) "A dynamic theory of organizational knowledge creation", *Organization Science*, Vol. 5, No. 1, pp 14-37.

Spradley, J.P. (1980) *Participant Observation*, Holt, Rinehart and Winston, New York.

Star, S.L. and Griesemer, J.R. (1989) "Institutional ecology, 'translations' and boundary objects: amateurs and professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39", *Social Studies of Science*, Vol. 19, No. 3, pp 387-420.

Trienekens, J.H., Wognum, P.M., Beulens, A.J.M. and Van der Vorst, J.G.A.J. (2012) "Transparency in complex dynamic food supply chains", *Advanced Engineering Informatics*, Vol. 26, No. 1, pp 55-65.

Wenger, E. (1999) *Communities of Practice: Learning, Meaning, and Identity*, Cambridge University Press, Cambridge.

How Does Innovation Activity Impact Firm Competitiveness? Evidence from Hungarian Firms

Anett Dobos, Zoltán Soma Kárpáti, Márk Recskó
Corvinus University of Budapest, Budapest, Hungary

Abstract

Innovation has long been recognized as a key driver of firm competitiveness; however, the mechanisms through which different innovation activities translate into performance remain insufficiently understood, particularly in Central and Eastern European contexts. While prior research has established a generally positive relationship between innovation and firm performance, empirical findings remain heterogeneous, and the underlying capability and knowledge-based mechanisms are not fully clarified. This paper investigates how innovation activities influence firm competitiveness through capability-based and knowledge-driven processes.

Building on the resource-based view (RBV) and the dynamic capabilities framework, the study extends existing approaches by incorporating a knowledge-based perspective, conceptualizing innovation as a process of knowledge creation, integration, and recombination. The empirical analysis is based on a sample of 332 Hungarian medium-sized and large firms. Using partial least squares structural equation modelling (PLS-SEM), the study examines the direct and indirect effects of process innovation, product and service innovation, and intellectual property ownership on firm competitiveness, measured by the Firm Competitiveness Index.

The results show that process innovation and intellectual property ownership positively influence competitiveness, primarily through operational excellence and organizational adaptivity. In contrast, product and service innovation activities exhibit a negative relationship with adaptivity and no significant positive effect on performance. These findings suggest that innovation does not translate into performance uniformly. Instead, its effects are mediated by the firm's ability to transform knowledge into capabilities.

From a knowledge-based perspective, process innovation enhances internal knowledge integration and efficiency, while intellectual property supports knowledge protection and recombination. Conversely, market-facing innovation may impose coordination and knowledge integration challenges that weaken adaptive capabilities, particularly in manufacturing-dominated contexts.

The study contributes to the IFKAD discourse by demonstrating that firm competitiveness is fundamentally capability-driven and rooted in knowledge asset management. The findings also have important managerial implications, suggesting that firms should sequence innovation investments by first strengthening internal capabilities before engaging in market-facing innovation activities.

Keywords: firm competitiveness, innovation activity, business performance, medium and large firms, Central Europe

Paper type: Academic Research Paper

1 Introduction

Innovation has been a central topic in economic and management research since Schumpeter (1934), who emphasized its role in driving economic development. Subsequent research has consistently linked innovation activities to firm performance and competitive advantage (Prajogo, 2006; Gunday et al., 2011). However, empirical findings remain heterogeneous, particularly across different institutional and regional contexts. In Central and Eastern Europe, empirical evidence on the innovation–performance relationship is relatively scarce. This gap is particularly relevant given the structural characteristics of these economies, where firms often operate under resource constraints and rely heavily on incremental innovation and capability development (Lukovszki et al., 2020).

This paper examines how innovation activities influence firm competitiveness through capability-based mechanisms. Building on the resource-based view (Barney, 1991) and the dynamic capabilities framework (Teece et al., 1997), the study adopts a knowledge-based perspective, interpreting innovation as a process of knowledge creation and recombination. In this framework, innovation does not directly lead to performance outcomes; rather, its impact is mediated through capabilities that enable firms to exploit and adapt their knowledge base. The paper focuses on three types of innovation activities: process innovation, product and service innovation, and intellectual property ownership. These are linked to two key capability dimensions: operational excellence and adaptivity, which together determine firm competitiveness.

The study contributes to the literature in three ways. First, it provides empirical evidence from a Central European context; second, it advances theoretical understanding by integrating innovation, capabilities, and knowledge assets into a unified framework. Third, it contributes to the IFKAD discourse by explicitly linking innovation activities to knowledge-based capability development.

2 Theoretical background

2.1 Innovation as a knowledge-based process

Innovation is commonly defined as the implementation of new or significantly improved products or processes (OECD, 2018). However, beyond this outcome-oriented definition, innovation can be interpreted as a knowledge-based process involving the creation, integration, and application of knowledge within the firm (Baregheh et al., 2009). From a knowledge-based view, firms achieve competitive advantage by effectively managing their knowledge assets (Grant, 1996). Innovation activities represent structured mechanisms through which knowledge is recombined and transformed into value. Different types of innovation reflect distinct knowledge processes. Process innovation enhances internal knowledge efficiency, product and service innovation relies on external and market knowledge, while intellectual property represents codified and protected knowledge.

2.2 Resource-based view, dynamic capabilities and knowledge assets

The resource-based view explains firm performance differences through resource heterogeneity and the possession of valuable, rare, inimitable, and organizationally embedded resources (Barney, 1991). However, in dynamic environments, static resources are insufficient to sustain competitive advantage. The dynamic capabilities framework extends the RBV by emphasizing the firm's ability to integrate, build, and reconfigure resources (Teece et al., 1997). Dynamic capabilities enable firms to respond to changing environments through sensing, seizing, and transforming activities (Teece, 2007). Integrating a knowledge-based perspective, capabilities can be interpreted as mechanisms through which knowledge is mobilized and transformed into performance outcomes. Operational excellence reflects the efficient exploitation of existing knowledge, while adaptivity represents the ability to explore and integrate new knowledge.

2.3 Innovation and firm competitiveness

A substantial body of empirical research demonstrates a positive relationship between innovation and firm performance (Prajogo, 2006; Gunday et al., 2011). However, this relationship is not uniform. The performance effects of innovation depend on how innovation activities are embedded within organizational processes and capabilities (Jin and Choi, 2019). Recent research increasingly emphasizes that innovation influences performance indirectly by shaping firm capabilities (Chikán et al., 2022). In this perspective, competitiveness is a multidimensional construct reflecting the combined effects of resources, capabilities, and strategic choices. Figure 1 illustrates the conceptual framework of the study, highlighting the direct and capability-mediated relationships between different types of innovation activities and firm competitiveness.

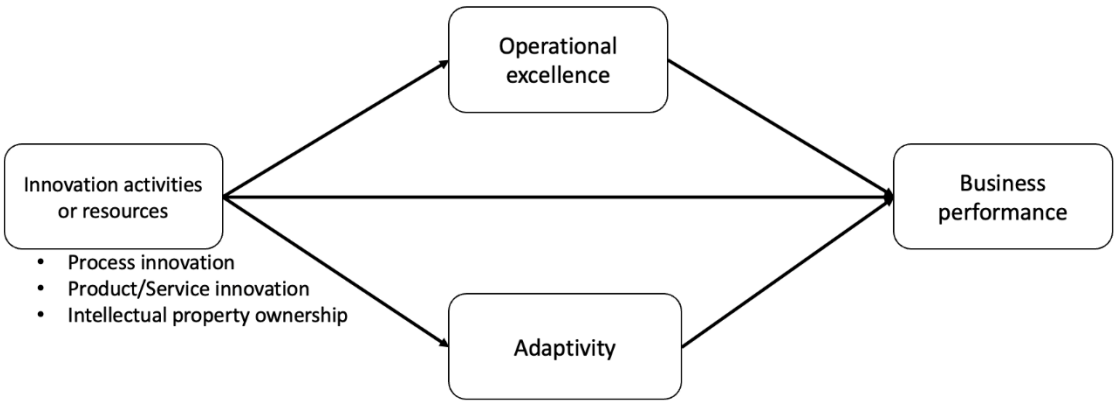


Figure 1: The relationship between innovation and firm performance.
Source: Author's own compilation

Based on the theoretical framework, the study examines the following hypotheses:

- *H1: Process innovation activities have both a significant direct and an indirect positive effect on firm competitiveness, with the indirect effect being mediated by operational excellence and adaptivity*
- *H2: Product/service innovation activities have both a significant direct and an indirect positive effect on firm competitiveness, with the indirect effect being mediated by operational excellence and adaptivity.*
- *H3: Intellectual property ownership (IPO) activities have both a significant direct and an indirect positive effect on firm competitiveness, with the indirect effect being mediated by operational excellence and adaptivity.*

3 Methodology

The empirical analysis is based on a survey conducted among Hungarian firms with more than 50 employees. The final sample consists of 332 firms across multiple sectors, including manufacturing, trade, and transportation. Firm competitiveness is measured using the Firm Competitiveness Index (Chikán et al., 2022), which captures both operational and dynamic capability dimensions. Operational excellence reflects efficiency and quality, while adaptivity captures the firm's ability to respond to environmental changes. Innovation activities are measured using composite indices constructed from binary indicators, following established methodological approaches (Eurostat, 2022). This allows capturing the breadth and intensity of innovation activities.

Table 1: Sample demographics. Source: Author's own compilation.

Characteristic	Description	Frequency	Percentage
Firm size	No. of employees in 2024		
	50-249	290	87.35%
	>250	42	12.65%
Ownership			
	Hungarian owned	301	90.66%
	Foreign owned	31	9.34%
	Family business	122	36.75%
Sectors			
	Manufacturing	162	48.80%
	Construction	33	9.94%
	Wholesale and retail trade; repair of motor vehicles	86	25.90%
	Transportation and storage	51	15.36%
Export activity			
	Export activity	202	60.84%
	No export activity	130	39.16%

Partial least squares structural equation modelling (PLS-SEM) is used to test the hypothesized relationships (Hair et al., 2022). This method is particularly suitable for exploratory models and complex relationships among latent constructs. To address potential robustness concerns, firm size and ownership structure are included as control variables. In addition, the analysis considers the potential influence of sectoral composition and reverse causality. While the cross-sectional nature of the data limits causal inference, the capability-mediated structure of the model provides theoretical support for the proposed relationships.

4 Results

The results indicate that operational excellence and adaptivity both have a significant positive effect on firm competitiveness (Figure 2). Process innovation exhibits a strong positive effect on both capability dimensions, which in turn translate into improved performance. Intellectual property ownership shows a significant positive effect on adaptivity, suggesting that codified knowledge assets enhance the firm's ability to respond to environmental changes. However, its direct effect on performance is not significant, indicating that its impact is primarily indirect. In contrast, product and service innovation activities do not have a significant positive effect on performance. Moreover, a negative relationship is observed between these activities and adaptivity, suggesting that market-facing innovation may impose coordination and integration challenges.

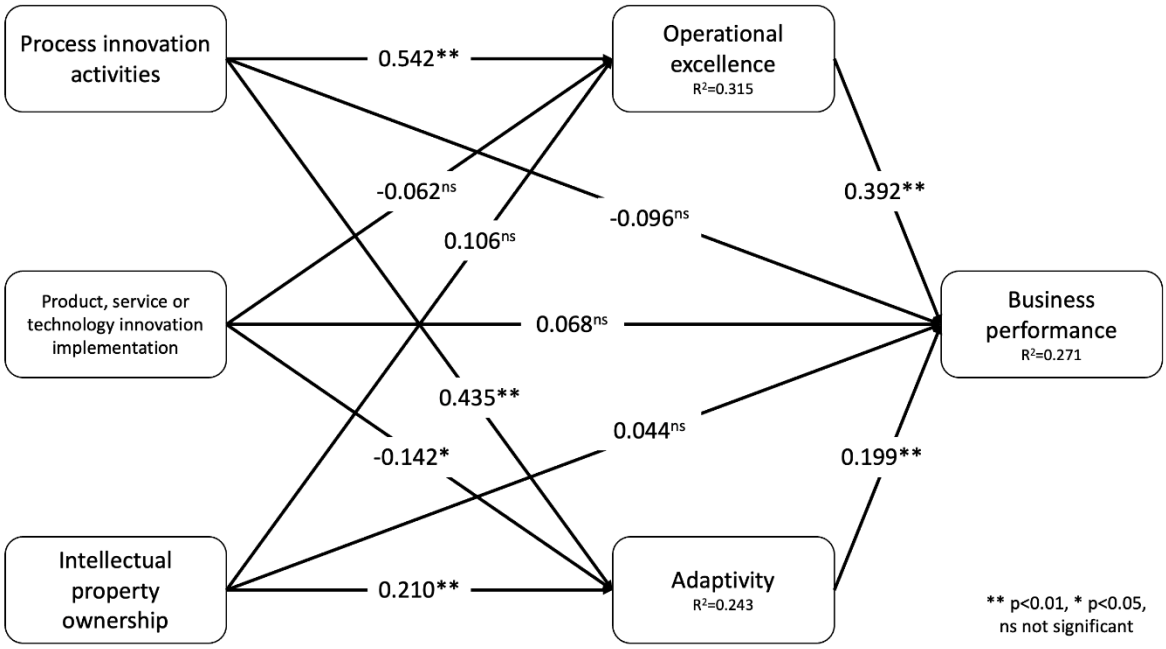


Figure 2: Assessment of the structural model.
Source: Author's own compilation.

5 Discussion

5.1 Results of the analysis

In this study, we explored how firms' innovation activities affect business performance both directly and indirectly through the mediating roles of operational excellence and organizational adaptivity. While the relationship between innovation and performance has been widely examined in the literature, empirical evidence from the Central and Eastern European region remains relatively scarce and insufficiently theorized. To contribute to this gap, we analysed data from 332 Hungarian firms and found that process innovation and intellectual property–related assets have a strong and positive influence on both operational excellence and adaptivity, thereby supporting hypotheses H1 and H3. However, our results did not support the second hypothesis. We observed a significant negative association between market-oriented innovation activities—such as product, service, and technology-related innovation—and adaptivity. At the same time, no statistically significant direct effect was identified between operational excellence and business performance. These findings indicate that different types of innovation do not contribute equally to capability development or performance outcomes, suggesting a more nuanced relationship between innovation and competitiveness.

The results are particularly meaningful in the Hungarian and broader Central European context, where both medium-sized and large firms account for a substantial share of employment and play a critical role in the economy (Hungarian Central Statistical Office, 2025). Our analysis underscores the importance of innovation activities, especially process innovation and intellectual property ownership, in shaping firm-level competitiveness. Considering that manufacturing firms make up a large proportion of the sample (over 48%) and that the Hungarian economy is relatively labor-intensive, it is both theoretically and economically reasonable that firms benefit more from improvements in internal processes and operational capabilities than from market-facing product or service innovation.

5.2 Theoretical contribution

The study contributes to the literature by integrating innovation, capabilities, and knowledge assets into a unified framework. It demonstrates that innovation does not directly drive performance; instead, its effects are mediated through capabilities that transform knowledge into value. This finding aligns with the dynamic capabilities perspective and extends it by emphasizing the role of knowledge integration and recombination. The results also contribute to the IFKAD discourse by highlighting the central role of knowledge-based capabilities in achieving firm competitiveness. Overall, the results support a capability-driven view of competitiveness, where innovation influences performance indirectly through capability development.

5.3 Managerial implications

Given that process innovation and intellectual property ownership were shown to have a significant impact on operational excellence, adaptivity, and overall business performance, these areas should be treated as strategic priorities by Hungarian firms. The combined development of process-oriented innovation and the acquisition or protection of intellectual property can contribute to achieving sustainable competitive advantage. Accordingly, managers need to deliberately invest in strengthening the underlying capabilities that support these activities.

Although innovation often involves substantial costs, the results suggest that such investments can generate long-term returns. To realize these benefits, firms must systematically build organizational and firm-level capabilities that enable continuous innovation and the development of dynamic capabilities. This is not an immediate outcome but rather a long-term process that requires sustained effort over several years, during which capability building itself can be interpreted as a strategic objective (Hutton, Demir & Eldridge, 2021). In practice, this implies that managers need to adopt a dynamic capabilities perspective, focusing on their ability to sense emerging opportunities, seize them, and reconfigure existing resources accordingly (Teece, 2016).

Organizations that perform strongly across these capability dimensions are more likely to generate higher-level and more transformative innovations (Ellonen, Wikström & Jantunen, 2009). Achieving this also requires firms to actively seek and utilize knowledge beyond their organizational boundaries. Establishing and maintaining collaborations with universities, research institutes, and other external partners can significantly enhance firms' innovation potential (Lin et al., 2020).

In addition, firms need to continuously integrate external knowledge into their operations, reflecting the importance of absorptive capacity within the knowledge-based view. This also depends on the firm's ability to attract and retain highly skilled professionals, particularly in research and development roles. In the Hungarian context, this represents a considerable challenge, as companies face ongoing labour shortages, outward migration of talent, and increasing competition for qualified employees in international markets. These findings reinforce the role of knowledge assets as the underlying drivers of capability development and firm competitiveness.

6 Conclusions

This study offers a comprehensive examination of how various types of innovation activities influence firm competitiveness through the mediating roles of operational excellence and organizational adaptivity. Based on a large sample of Hungarian medium-sized and large enterprises, the results indicate that the relationship between innovation and business performance is neither direct nor uniform. Rather, the impact of innovation is primarily indirect, unfolding through the development of internal capabilities related to efficiency and adaptability.

From an empirical perspective, the findings emphasize the importance of process innovation and intellectual property–related resources in enhancing both operational excellence and adaptivity, thereby contributing to higher levels of firm competitiveness. These results support the assumptions of the resource-based view and the dynamic capabilities framework, demonstrating that sustained performance advantages arise not only from innovation outcomes themselves but from the firm's ability to translate innovation activities into effective organizational capabilities. In contrast, innovation activities that are oriented toward markets—such as product, service, and technology innovation—do not necessarily lead to improved performance and may, in some cases, place additional pressure on adaptive capacities, highlighting potential trade-offs in firms' innovation strategies.

Furthermore, the analysis reveals that the relationships identified between innovation, capabilities, and performance remain consistent across different firm sizes and ownership structures. Neither firm size nor foreign ownership significantly modifies these core relationships, suggesting that the underlying mechanisms operate similarly across organizational contexts. This consistency enhances the robustness and broader applicability of the proposed model within the Hungarian economy and, more generally, in the Central and Eastern European region.

7 Ethics Declaration

The research reported in this paper was conducted with prior ethical approval granted by the Research Ethics Committee of Corvinus University of Budapest. The approval was issued under file number KRH/181/2025 for the research project "Competitiveness Research Center – Firm competitiveness research 2025–2027".

The study is based on survey data collected from company representatives. Participation in the research was voluntary, and informed consent was obtained from all respondents prior to data collection. Respondents were

contacted by a professional data collection provider, and they were fully informed about the purpose of the research, the voluntary nature of participation, and their right to withdraw at any time. All data were collected and processed in accordance with applicable data protection regulations. The dataset was anonymized prior to analysis, and no individual firm or respondent can be identified. Researchers had access only to aggregated and anonymized data, ensuring confidentiality and compliance with ethical standards.

The research did not involve vulnerable groups, sensitive personal data, or any procedures that could pose physical, psychological, or social risks to participants. The study adhered to the ethical principles and institutional guidelines of Corvinus University of Budapest throughout the research process.

8 AI Declaration

The authors declare that Grammarly was used during the preparation of this manuscript for language checking and to improve grammar, clarity, and readability.

The use of this tool was limited to linguistic refinement only. All substantive elements of the research, including the research design, data analysis, interpretation, and conclusions, were developed solely by the authors, who take full responsibility for the content of the paper. No AI tools were used for data analysis, empirical modelling, or the generation of research findings.

References

Baregheh, A., Rowley, J. and Sambrook, S. (2009) 'Towards a multidisciplinary definition of innovation', *Management Decision*, 47(8), pp. 1323–1339.

Barney, J. (1991) 'Firm resources and sustained competitive advantage', *Journal of Management*, 17(1), pp. 99–120.

Chikán, A., Czakó, E., Kiss-Dobronyi, B. and Losonci, D. (2022) 'Firm competitiveness: a general model and a manufacturing application', *International Journal of Production Economics*, 243, 108316.

Ellonen, H.K., Wikström, P. and Jantunen, A. (2009) 'Linking dynamic-capability portfolios and innovation outcomes', *Technovation*, 29(11), pp. 753–762.

Eurostat (2022) *Oslo Manual guidelines*. Luxembourg: Publications Office of the European Union.

Grant, R.M. (1996) 'Toward a knowledge-based theory of the firm', *Strategic Management Journal*, 17(S2), pp. 109–122.

Gunday, G., Ulusoy, G., Kilic, K. and Alpkan, L. (2011) 'Effects of innovation types on firm performance', *International Journal of Production Economics*, 133(2), pp. 662–676.

Hair, J.F., Hult, G.T.M., Ringle, C.M. and Sarstedt, M. (2022) *A primer on partial least squares structural equation modeling (PLS-SEM)*. 3rd edn. Thousand Oaks: Sage.

Hungarian Central Statistical Office (2025). *Performance indicators for businesses by small and medium-sized enterprise category*. Available at: https://www.ksh.hu/stadat_files/gsz/hu/gsz0018.html

Hutton, D., Demir, R. and Eldridge, S. (2021) 'The role of dynamic capabilities in strategy', *Long Range Planning*, 54(5).

Jin, S.H. and Choi, S.O. (2019) 'The effect of innovation capability on business performance', *Sustainability*, 11(19), 5246.

Lin, H., Zeng, S., Liu, H. and Li, C. (2020) 'How do intermediaries drive corporate innovation?', *Technological Forecasting and Social Change*, 161.

Lukovszki, L., Rideg, A. and Sipos, N. (2020) 'Resource-based view of innovation activity in SMEs', *Journal of Entrepreneurship*, 29(1), pp. 77–96.

OECD (2018) *Oslo Manual 2018: Guidelines for collecting, reporting and using data on innovation*. Paris: OECD Publishing.

Prajogo, D.I. (2006) 'The relationship between innovation and business performance', *Knowledge and Process Management*, 13(3), pp. 218–225.

Schumpeter, J.A. (1934) *The theory of economic development*. Cambridge, MA: Harvard University Press.

Teece, D.J. (2007) 'Explicating dynamic capabilities', *Strategic Management Journal*, 28(13), pp. 1319–1350.

Teece, D.J., Pisano, G. and Shuen, A. (1997) 'Dynamic capabilities and strategic management', *Strategic Management Journal*, 18(7), pp. 509–533.

Accounting Standardization and Circular Knowledge Frictions in Cooperative Banks

Andrea Comandé, Olga Ferraro
University of Calabria, Italy

Abstract

This research investigates the organizational and strategic transformation of Cooperative Credit Banks (CCBs) under the pressure of increasing regulatory standardization. In the contemporary European banking landscape, the transition toward hard data metrics (e.g., Basel Regulation, EBA guidelines) poses a significant challenge to the traditional business model of CCBs, which is historically rooted in 'Relational Capital' and the processing of 'soft information'. The study aims to analyse how this exogenous regulatory shock acts as a catalyst for internal reconfiguration, shifting these hybrid organizations from a model of 'Circular Knowledge' toward one of 'Linear Efficiency'. The paper adopts an inductive qualitative methodology based on a multiple-case study design. The proposed research setting is enabled by a strategic partnership with a pilot CCB, complemented by a comparative analysis of additional cases selected through specific control. Drawing on Transaction Cost Economics (TCE) as the primary theoretical lens, the study employs semi-structured interviews with top management and internal document analysis to map the shift in decision-making authority and credit-granting processes. As a research-in-progress, this study proposes a theoretical framework to be further validated through the ongoing empirical phase. By moving beyond a purely normative view of regulation, this study aims to contribute to Knowledge Management and Organization Studies by illustrating how measurement frictions can structurally alter the boundaries and the identity of hybrid financial institutions. The findings will provide both scholars and practitioners with a critical understanding of the unintended organizational costs of systemic banking supervision.

Keywords: Cooperative Credit Banks, Relational Capital, Circular Knowledge, Banking Biodiversity

Paper type: Academic Research Paper

1 Introduction

Hybrid organizations, such as Cooperative Credit Banks (CCBs), represent a distinct organizational paradigm characterized by a dual logic: pursuing economic objectives while fulfilling a social mission (Bruni and Zamagni, 2004; Ferraro et al., 2024). Following an evolutionary economic perspective (Williamson, 1985), markets tend to select organizational forms that minimizes transaction and ownership costs most efficiently. While institutional factors – such as regulatory protections, tax exemptions, or public subsidies - can distort this mechanism, the historical persistence of CCBs must be explainable by a specific relative advantage that justifies their existence in purely economic terms. Particularly, we argue that the core of this advantage lies in what this study defines as Relational Capital. We specialize this definition conceptualizing it as a strategic informational asset generated through an iterative learning process wherein qualitative data extracted by means of repeated, localized interactions materialize into implicit contracts and reputational constraints, thereby enabling the design of optimal contracts and enforcement schemes.

By circumscribing their activities to localized areas, these institutions specialize in high frequency, lower complexity financial transactions. Within this framework, relational capital functions as a primary informational resource, as it addresses information asymmetries through mechanisms that fundamentally differ from market signals (Scott et al., 2001). The accumulation over time of knowledge and uncodified data (soft information), enabled by the embedded nature of these organizations, defines a distinct specialization model whereby relational proximity and repeated interaction efficiently substitute formal quantitative metrics in assessing creditworthiness. In this context, the CCB functions as a node in a circular system of knowledge: local savings are reinvested locally, fuelled by a continuous feedback loop of soft information that enables the efficient mitigation of information asymmetries, continuously generating knowledge that create the condition for this specific organizational configuration to optimally operate in the market.

However, the banking institutional environment in Europe is marked by a pervasive process of regulatory standardization driven by a calculative logic intended to minimize monitoring and agency cost in the European financial system. By means of standardized accounting metrics (such as Basel Regulation and the EBA NPL guidelines) regulators establish a formal evaluative scaffolding to reduce informational gaps at the systemic level. This process mirrors what has been described in sociological literature as 'commensuration' (Espeland and Stevens, 1998): the synthesis of highly specific, relationship-reducing mechanisms into defined categories

to ensure visibility and comparability at the expense of the relevance of the local context. That is, displacing soft information – although fundamental for an accurate assessment of a prospective borrower's true risk profile – prioritizing hard information, as the latter can be seamlessly integrated into a uniform and objectively verifiable supervisory system (Behn et al., 2016). In doing so, these mechanics create a formal evaluative foundation to manage uncertainty and incomplete informational sets, reflecting the regulator's need to stabilize decision-making processes and thoroughly monitor risks across the financial system.

We argue that by imposing standardized metrics onto relationship governed by implicit contracts, the regulatory framework introduces steep measurement cost over CCBs. In particular, our argument posits that an approach failing to account for the idiosyncratic nature of local lending might disrupt the circularity of information among delimited boundaries, and in turn this may lead CCBs to abandon their specialized informational technology, reassessing their perceived mission in the market and ultimately their behaviour. Centralized accounting regimes significantly restrict managerial discretion by displacing relational logics, ultimately emerging as the sole legitimate foundation for defining and measuring variables relevant to investment decisions. The evaluative scaffolding systematically excludes those intangible assets, remaining outside the visibility of conventional reporting framework.

Sociological and Critical Accounting perspectives provide valuable complementary insights. Indeed, when regulators impose standardized metrics to minimize their own monitoring costs, accounting acts as a powerful mechanism of commensuration (Espeland and Stevens, 1998). In this view, accounting is not merely a neutral, technical recording of economic facts, but an institutional practice that shapes organizational reality by altering what is rendered visible to decision makers (Hopwood, 1987). Indeed, once established, these standardized metrics tend to acquire a self-justifying force (Espeland and Sauder, 2007), where the necessity to continuously evaluate organizational outcomes against them drives a shift in organizational behaviour, promoting homogenization and displacing other types of authority and judging systems (Mennicken and Espeland, 2019), ultimately redefining what action can be considered legitimate. The magnitude of this progressive alignment will depend on the intensity of the coercive pressure and the critical need to comply with these metrics to preserve operational legitimacy (Meyer and Rowan, 1977). The basic idea of this research, that will inform the formulation of the propositions, is that when managerial discretion is displaced by rigid decision-making processes, the resources invested in cultivating Relational Capital are no longer capitalized.

Through a critical analysis of these valuative frictions, this paper intend to investigate over the frictional costs of regulatory homogenization for decentralized financial ecosystems. Particularly, whilst broader macro-trends, such as digitalization, concurrently affect bank-customer interactions, this paper focuses strictly on the organizational impact of these accounting and regulatory pressures. Methodologically, the study adopts a theory-building approach grounded in case study research (Eisenhardt, 1989). Within this framework, our research objective is twofold; that is: *i)* investigating how standardized accounting alters the regime of organizational visibility; *ii)* examining the consequences of isomorphic dynamics in these contexts, informing an initial set of propositions regarding the potential for "mission drift" (Battilana and Dorado, 2010).

Following Eisenhardt's iterative methodology, the case study will not be employed merely to validate these *a priori* hypotheses. Rather, it will serve as an empirical setting to test and continuously refine the theoretical framework, allowing the initial constructs to be modified or expanded in response to the specific organizational dynamics that emerge from the data.

2 The Architecture of Soft Information: Local Knowledge and the Cooperative Advantage

In localized settings, economic transactions are rarely isolated events; rather, they unfold within a dense network of repeated interactions among a defined set of actors. Because information about past behaviour is shared in local boundaries, historical interactions and mechanisms of reciprocity function as credible commitments. Consequently, opportunistic behaviour is deterred by the implicit threat of social ostracism and reputational damage, thereby facilitating sustainable, long-term economic cooperation (Coleman, 1988). The proximity of the relationship may, for instance, translate into greater contractual flexibility, allowing for discretion and renegotiation according to particular circumstances; more effective monitoring of collateral and contractual covenants, or an expansion of credit provision in cases where the rents derived from the expected value of the relationship compensate for the costs of financing projects that are initially unprofitable (Boot, 2000).

In support of our argument, we align with Hayek's identification of the primary economic problem: the adaptation to localized "circumstances of time and space". Although Hayek's concern was primarily epistemological, he underlined the need for recognizing the complexity in idiosyncratic, local knowledge, that by its very nature resists aggregation into standardized statistical measures and yet constitutes the essential

basis for economic adaptation within specific contexts (Hayek, 1945, 1967; Williamson, 1985). While literature has widely documented the advantages of 'relationship lending' dynamics in terms of enhanced resilience and risk management in financial markets (Petersen and Rajan, 1994; Boot, 2000; Ferri et al., 2014), the deep reliance on proprietary soft information requires specific organizational architectures for the continuous accumulation of uncodified data (soft information), defining a distinct specialization path.

For instance, Stein (2002) argues that relatively flat organizational structures are more efficient when dependent upon *soft information* (or equivalently, if the organization is specialized in soft data collection). This perspective helps explain the historical resistance of German and Italian Cooperative Banks to the centralizing pressures of the European Central Bank toward centralization. The resulting compromise - like the aggregation within comprehensive groups in Italy with sub-banks who own the group and retain varying degree of autonomy - internalized the tension transforming the central holding into a transmission mechanism for regulatory pressures, allowing for a general reduction of Monitoring costs for the ECB.

In vast, complex, and interconnected markets, transaction costs are minimized through formal institutional rules that are necessary to manage information asymmetry and to constrain opportunistic behaviour among distant agents where defining and enforcing agreements is costly (North, 1991). For example, Agarwal et al. (2014), argue that centralized regulators are bound to remain detached from local socio-economic dynamics, adopting a generalized evaluative posture that ensures systemic consistency and significantly lower economic costs. In particular, local supervisory forbearance - often driven by the intent to shield regional economies - correlates with delayed recognition of problem loans and higher bank failure rates.

However, albeit relationship lending in closed, auto-referential settings theoretically exposes the institution to severe agency costs, cronyism, rent-seeking and exposes borrowers to hold-up problems (Sharpe, 1990; Rajan, 1992; Uzzi, 1997; Weinstein and Yafeh, 1998), the cooperative architecture is functional to the mitigation of these dynamics. In particular, micro-prudential regulation has been devised to solve specific agency problem inherent to commercial banking, to prevent the moral hazard stemming from limited liability and institutional safety nets, by means of rigid capital requirements and standardized (risk-weighted) metrics (Freixas and Rochet, 2008). In the case of CCBs, the 'common bond' (Hansmann, 1996) - such as territorial or associative proximity - reduces adverse selection and moral hazard through incentive alignment that constrain opportunistic behaviour drawing upon social pressures, the credible threat of reputational loss within the business community, and informal peer sanctions (Stiglitz, 1990; Banerjee et al., 1994). For the analytical purpose of this study, we assume this organizational form as our theoretical baseline to isolate the specific frictions generated by exogenous regulatory commensuration.

2.1 Regulatory standardization as an exogenous shock

Following from the Hansmann (1996) arguments, the persistence of the cooperative form can be viewed as an efficient organizational response to minimize transaction costs associated with locally rooted information asymmetries. However, regulatory homogenization through the widespread implementation of impersonal logics onto CCBs generates two structural frictions, particularly: *i)* replacing soft information with metrics based on hard data prevents the assessment of physiologically "opaque" subjects (e.g., micro-enterprises, artisans) thereby leading to systemic credit rationing for local borrowers; and at a more aggregate level *ii)* the reduction of "banking biodiversity" by imposition of isomorphic logics.

The methodological architecture of traditional financial accounting operates on a linear, transactional logic. That is, it excels at measuring standardized, hard data, but inherently lacks the vocabulary to capture the effects of facts pertaining to the creation and maintenance of relational capital into commensurable metrics. This limitation stems, obviously, from the structural difficulty of codifying phenomena that are predominantly qualitative in nature, highly specific to a certain relationship or network and poorly transferable (Petersen and Rajan, 1994). By collapsing vast amounts of idiosyncratic information into standardized metrics, this process drastically diminishes the relevance of local context, rendering the very definitions of what is considered legitimate and rational strictly dependent on what can be quantitatively measured (and maximized), anchoring decision-making to an institutional discourse reframing the foundations upon which what is considered "optimal" is established.

The exogenous imposition through institutional pressures also carries a dominant societal discourse regarding what constitutes a legitimate organization (Meyer and Rowan, 1977). As the domain of what is visible shifts to comply with these exogenous rules, it directly alters the incentive structure guiding managerial behaviour (Mennicken and Espeland, 2019). For example, it becomes costly to produce knowledge that cannot be credibly transmitted along the hierarchy, anticipating that the effort will be wasted since it will not inform standardized capital allocation procedures (Stein, 2002). As organizations adhere to and start to privilege

standardized metrics, the immediate costs of investing in relational capital will appear more certain in the relative outcome, with more specific effects distributed in shorter time horizons. The proximity and clarity of such causalities, entails consequences for how resources are allocated across space and time. Furthermore, increasing salience of measurable attributes of transaction constrains managerial discretion, since decision making rely on information that can be formally validated, reducing the role of tacit knowledge and local judgement.

As these changes becomes the basis for the activities carried out by the organization, structures tend to adapt following resulting changes in transaction (that is, resource allocation) to align with the dominant evaluative framework. This may include increased centralization, formalization of procedures and a shift toward transactional forms of interaction. Consequently, time and resources invested by organizations - such as CCBs in cultivating soft information to maintain a capillary territorial presence, engaging in continuous face-to-face qualitative listening – are not capitalized as strategic assets.

The friction between what is measurable and what remains invisible within accounting systems directly intersects with the challenge of quantifying non-financial ESG outcomes. Insofar as we conceptualize the 'social dimension' of CCBs - traditionally defined by mutuality and territorial ties – as a circular knowledge infrastructure, in this framework tacit knowledge extracted from the local environment is analysed and reinvested within the same boundaries, creating a closed-loop feedback systems that fuels local sustainable investment.

2.2 The friction between local knowledge and systemic metrics

Historically, particularly within the European Context (such as in Italy and Germany), cooperative banks emerged as a local response to community needs, as a mean for financial inclusion and emancipation from local monopolistic practices, such as usurious interest rates. Over decades, the continuous refinement of processes tailored to these specific local needs sedimented a distinct organizational culture which is fundamentally misaligned with purely profit-maximizing logics of commercial banking in competitive settings. Oftentimes, this kind of institution tends to promote themselves as agents fundamentally different from traditional banks, tied to the local community in which they operate and to a general set of ethical value, such as the ones defined by the International Cooperative Alliance (ICA), in which they recognizes themselves and that ideally should characterize their daily operations, governance and stakeholder relations.

However, the macro-institutional environment has historically shifted to structurally alter the competitive landscape in which CCBs operate. The introduction of modern banking regulations and deposit insurance schemes (e.g., FITD in Europe or FDIC in the US) has artificially lowered the contracting and transaction costs for depositors of traditional commercial banks by providing a systemic safety net, implicitly subsidizing the moral hazard inherent in the commercial model, and ultimately harming the competitive advantage that historically attracted depositors to cooperative institutions (Hansmann, 1996). The imposition of standardized regulatory regimes can be conceptualized as institutional interventions that potentially reconfigure the transaction cost structure of credit intermediation, introducing significant informational frictions when applied to relational lending models by altering the measurement and monitoring costs. In turn, changes to the nature and characteristics of transactions follows changes in governance that should efficiently minimize transaction costs (Williamson, 1985). Again, this logical sequence might explain the organizational transformations across Europe of Cooperative Credit sector following the imposition of new ECB/EBA's requirements.

By tracing a parallel with the idea of social structure premised in the Economic Sociology, Uzzi (1997) acknowledges that organizations that build their competitive advantage upon embedded ties face a substantial risk of failure when institutional changes rationalize the basis of, or preclude the formation of new ones. To sustain these arguments, we can recur to two examples. Firstly, the evaluative structure of supervisory frameworks structurally privileges hard information (e.g., audited financial statements, standardized collateral) (Power, 2004). Consequently, credit relationship that are implicitly sustained by continuous interaction history and mutual trust may be interpreted as 'opaque' within this framework, thereby assigning higher Risk-Weighted Asset coefficients, and forcing CCBs to set aside greater amounts of costly regulatory capital with respect to a theoretical risk profile that may diverge from the actual, artificially driving up the cost of local credit allocation. For instance, Berger and Udell (2006) find that, although there are transactional technologies specifically designed to evaluate and serve this kind of clients by recurring to hard data diverse from financial statements, rigid rules and aggregative logics as for Basel II implicitly encourage the adoption of such technologies at the expense of relationship lending based on soft information. We can also pose a symmetrical reasoning where this dynamic operate counterfactually as a constraint: credit may simply be rationed and not granted at all, as the regulatory capital requirements disincentivize the transaction from the outset (Peek and Rosengren, 1995; Aiyar

et al., 2014). What is important here is that, in more general terms, to changes in operational target correspond modifications on the feasible set of credit allocation choices available within the banking sector. Secondly, recent regulatory approaches to Non-performing Loans such as the EBA's calendar provisioning mandates, impose automatic, time-bound write-downs of impaired credit. This mechanical accounting treatment disregards the flexibility and localized recovery strategies that CCBs historically employed to manage borrowers facing temporary liquidity shocks. By forcing premature loss recognition, the regulation actively disincentivizes lending to economic actors who may exhibit irregular cash flows despite the possibility of private information regarding the actual long-term solvency. The specific value of idiosyncratic knowledge is depreciated in front of the imposition of automatic write-downs and forced asset sales (Moretto and Parigi, 2025) or by disregarding the role of managerial discretion in the forward-looking recognition of credit risk (Bushman and Williams, 2012).

Although these metrics are introduced with the legitimate goal of reducing systemic risk and ensuring transparency across financial systems, they are designed for contexts of impersonal exchange as a mean for creating public, impartial knowledge (Porter, 1995). The issue can be most succinctly summarized by the idea that, because regulatory apparatus discourse lacks the vocabulary to measure the value of soft information in these settings, it prudentially assumes that such information does not exist other than in its indirect consequences, and the ultimate outcome is that of reintroducing one of the primary frictions the cooperative model emerged to cope with.

3 Proposed Research Design

Whilst it is beyond the scope of this paper to normatively assess the social optimality of these mechanics, our research objective is strictly positive. We aim to empirically investigate whether and how exogeneous increase in measurement and compliance costs for the CCBs, may entail internal rearrangements with respect to the specificity of transactions, that we contend may be different in nature and scope because of the resulting changes in the incentives structure that ultimately defines and constrain economic behaviour in the market.

3.1 Empirical strategy and data sources

Methodologically, this study benefits from an immersive research setting facilitated by a formal research partnership with a prominent Cooperative Credit Bank. While such proximity necessitates rigorous analytical protocols to mitigate researcher bias, it grants unprecedented access to multiple organizational layers and rich data collection to build inductive theory. To ensure systematic data triangulation, the empirical investigation relies upon qualitative organizational routines and institutional outputs. Namely:

- *Qualitative Data*: the inductive process relies on field observations and semi-structured interviews with top management, risk officers and local branch managers. This ethnographic immersion is corroborated by access to archival data, such as internal credit committee minutes (where available), allowing us to trace how soft information is debated and utilized in credit allocation policies, tracing decision-making process from policy-setting by the Board of Directors to execution by Executive Committees.
- *Secondary Archival Data*: to contextualize and verify the consistency of narration, interview data will be cross-referenced with institutional documentation. This includes the analysis of group-level credit policies, annual mission reports and aggregated financial statements.

Ideally, the juxtaposition of these sources is intended to neutralize idiosyncratic biases, allowing for the isolation of governance rearrangements following exogenous regulatory shock.

3.2 Analytical focus and theoretical sampling

We focus our empirical setting on the Italian Cooperative Credit System (Banche di Credito Cooperativo), adopting a multiple-case study approach that draws on Eisenhardt's (1989) protocol. That is, we employ theoretical sampling to isolate the effect of the regulatory shock from confounding factors. In particular, the selected CCBs are controlled along the following dimensions:

- *Size (Total Assets)*: to ensure that organizational changes are not the result of scale economies (Berger and Udell, 2005), the selected institutions operate within a similar dimensional threshold.
- *Historical NPL Ratio (as a proxy of pre-shock health)*: to avoid attributing organizational rearrangements to restructuring operations or excessive focus on costs saving (Berger and DeYoung, 1997). The sample focuses on CCBs that exhibit sound financial health over time.
- *Macroeconomic Context:* the sample focuses on economic institutions operating within comparable socio-economic environments, aiming at neutralizing variations due to local credit demand fluctuations (Guiso et al., 2004).

- *Leadership Stability:* since modifications in the organizational governance may be due to changes in the Top Management (Cuñat and Garicano, 2010).
- *Banking Group Affiliation:* to control for internal policies the selected cases belong to the same Cooperative Banking Group.

4 Expected Propositions and Implications

This study develops a set of analytical propositions to guide the empirical investigation. They are not intended as deterministic predictions, but as theoretically informed expectations subject to refinement through inductive analysis. In particular:

- *Internal Centralization and Bureaucratization*: following Stein (2002), regulatory mandates for standard metrics to flatten the comparative advantage of decentralized structures, leading to an elongation of the internal hierarchies and counter-balancing discretionary power of local branch managers with central risk and compliance departments.
- *Measurement Frictions and Cost-Income Dynamics:* the mandatory translation of idiosyncratic soft information into standardized regulatory categories acts as a severe transactional friction. This artificially inflates measurement and compliance costs, driving a structural upward pressure on the Cost-Income ratio of CCBs.
- *Transactional Shift*: as the internal cost of processing soft information rises, CCBs face structural incentives to crowd-out relationship lending, progressively shifting their credit portfolios toward standardized, transactional financial products.
- *Depreciation of the "Common Bond"*: we aim to observe how internal governance mechanisms autonomously adapt beyond formal compliance. Particularly, we expect that peer-monitoring and reputational constraints (Hansmann, 1996; Stiglitz, 1990) intrinsic to bounded settings - such as a cooperative community – are progressively marginalized in formal credit approval processes.

These dynamics may suggest that regulatory commensuration may reshape not only operational processes but also the informational foundations of cooperative banking. This raises a potential trade-off between systemic comparability and the preservation of locally embedded informational advantages, with implications for both regulators and practitioners.

5 Ethics Declaration

This study does not require an Ethics Committee Approval.

6 AI Declaration

During the preparation of this work the authors used Gemini in order to improve language and clarity of the manuscript. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

References

Agarwal, S., Lucca, D., Seru, A. and Trebbi, F. (2014) "Inconsistent regulators: Evidence from the banking industry", The Quarterly Journal of Economics, Vol. 129, No.2, pp. 889–938.

Aiyar, S., Calomiris, C. W. and Wieladek, T. (2014) "Does macro-prudential regulation leak? Evidence from a UK policy experiment", Journal of Money, Credit and Banking, Vol. 48, No. 1, pp. 181–214.

Banerjee, A. V., Besley, T. and Guinnane, T. W. (1994) "Thy neighbor's keeper: The design of a credit cooperative with theory and a test", The Quarterly Journal of Economics, Vol. 109, No. 2, pp. 491–515.

Battilana, J. and Dorado, S. (2010) "Building sustainable hybrid organizations: The case of commercial microfinance organizations", Academy of Management Journal, Vol. 53, No. 6, pp. 1419–1440.

Behn, M., Haselmann, R. and Vig, V. (2022) "The limits of model-based regulation", The Journal of Finance, Vol. 77, No. 3, pp. 1635–1684.

Berger, A. N. and DeYoung R. (1997) "Problem loans and cost efficiency in commercial banks", Journal of Banking & Finance, Vol. 21, No. 6, pp. 849–870.

Berger, A. N., Miller, N. H., Petersen, M. A., Rajan, R. G. and Stein, J. C. (2005) "Does function follow organizational form? Evidence from the lending practices of large and small banks", Journal of Financial Economics, Vol. 76, No. 2, pp. 237–269.

Berger, A. N. and Udell, G. F. (2006) "A more complete conceptual framework for SME finance", Journal of Banking & Finance, Vol. 30, No. 11, pp. 2945–2966.

Boot, A. W. A. (2000) “Relationship banking: What do we know?”, Journal of Financial Intermediation, Vol. 9, No. 1, pp. 7–25.
Bruni, L. and Zamagni, S. (2004) Economia civile: Efficienza, equità, felicità pubblica. Il Mulino.
Bushman, R. M. and Williams, C. D. (2012) “Accounting discretion, loan loss provisioning, and discipline of banks’ risk-taking”, Journal of Accounting and Economics, Vol. 54, No. 1, pp. 1–18.
Coleman, J. S. (1988) “Social capital in the creation of human capital”, American Journal of Sociology, Vol. 94, pp. S95–S120.
Cuñat, V. and Garicano, L. (2010) “Did Good Cajas Extend Bad Loans? Governance, Human Capital, and Loan Portfolios”, FEDEA working paper series (2010-08).
Eisenhardt, K. M. (1989) “Building theories from case study research”, The Academy of Management Review, Vol. 14, No. 4, pp. 532–550.
Espeland, W. N. and Sauder, M. (2007) “Rankings and reactivity: How public measures recreate social worlds”, American Journal of Sociology, Vol. 113, No. 1, pp. 1–40.
Espeland, W. N. and Stevens, M. L. (1998) “Commensuration as a social process”, Annual Review of Sociology, Vol. 24, pp. 313–343.
Ferraro, O., Cristiano, E. and Veltri, S. (2024) “The quality of voluntary sustainability reports in the Italian cooperative credit banks”, Journal of Co-operative Organization and Management, Vol. 12, No. 1.
Ferri, G., Kalmi, P. and Kerola, E. (2014) “Organizational structure and exposure to crisis among European banks: Evidence from rating changes”, Journal of Entrepreneurial and Organizational Diversity, Vol. 3, No. 1, pp. 35–55.
Frame, W. S., Karels, G. V. and McClatchey, C. (2002) “The effect of the common bond and membership expansion on credit union risk”, The Financial Review, Vol. 37, pp. 613–636.
Freixas, X. and Rochet, J.-C. (2023) Microeconomics of banking. 3rd ed. MIT Press.
Guiso, L., Sapienza, P. and Zingales, L. (2004) “The role of social capital in financial development”, American Economic Review, Vol. 94, No. 3, pp. 526–556.
Hansmann, H. (1996) The ownership of enterprise. Harvard University Press.
Hayek, F. A. (1945) “The Use of Knowledge in Society”, The American Economic Review, Vol. 35, No. 4, pp. 519–530.
Hayek, F. A. (1967) Studies in Philosophy, Politics and Economics. London: Routledge & Kagan Paul.
Hopwood, A. G. (1987) “The archaeology of accounting systems”, Accounting, Organizations and Society, Vol. 12, No. 3, pp. 207–234.
Mennicken, A. and Espeland, W. N. (2019) “What’s new with numbers? Sociological approaches to the study of quantification”, Annual Review of Sociology, Vol. 45, pp. 223–245.
Meyer, J. W. and Rowan, B. (1977) “Institutionalized organizations: Formal structure as myth and ceremony”, American Journal of Sociology, Vol. 83, No. 2, pp. 340–363.
Moretto, M. and Parigi, B. M. (2025) “The impact of calendar provisioning on bank strategies”, International Review of Economics & Finance, Vol. 101.
North, D. C. (1991) “Institutions”, Journal of Economic Perspectives, Vol. 5, No. 1, pp. 97–112.
Peek, J. and Rosengren, E. (1995) “The capital crunch: Neither a borrower nor a lender be”, Journal of Money, Credit and Banking, Vol. 27, No. 3, pp. 625–638.
Petersen, M. A. and Rajan, R. G. (1994) “The benefits of lending relationships: Evidence from small business data”, The Journal of Finance, Vol. 49, No. 1, pp. 3–37.
Porter, T. M. (1995). Trust in numbers: The pursuit of objectivity in science and public life. Princeton, NJ: Princeton University Press.
Power, M. (2004) The risk management of everything: Rethinking the politics of uncertainty. Demos.
Rajan, R. G. (1992) “Insiders and outsiders: The choice between informed and arm’s-length debt”, The Journal of Finance, Vol. 47, No. 4, pp. 1367–1400.
Scott Frame, W., Karels V. and McClatchey C., 2002. "The Effect of the Common Bond and Membership Expansion on Credit Union Risk," The Financial Review, Eastern Finance Association, vol. 37, No. 4, pp. 613-636.
Sharpe, S. A. (1990) “Asymmetric information, bank lending, and implicit contracts: A stylized model of customer relationships”, The Journal of Finance, Vol. 45, No. 4, pp. 1069–1087.
Stein, J. C. (2002) “Information production and capital allocation: Decentralized versus hierarchical firms”, The Journal of Finance, Vol. 57, No. 5, pp. 1891–1921.
Stiglitz, J. E. (1990) “Peer monitoring and credit markets”, The World Bank Economic Review, Vol. 4, No. 3, pp. 351–366.
Uzzi, B. (1997) “Social structure and competition in interfirm networks: The paradox of embeddedness”, Administrative Science Quarterly, Vol. 42, No. 1, pp. 35–67.
Weinstein, D. E., and Yafeh Y. (1998) “On the Costs of a Bank-Centered Financial System: Evidence from the Changing Main Bank Relations in Japan” The Journal of Finance, Vol. 53, No. 2, pp. 635–72.
Williamson, O. E. (1985) The economic institutions of capitalism. Free Press.

Intelligent Knowledge Management Systems for Human-AI Collaboration in Financial Decision-Making: A Behavioural Finance Perspective on Value Creation Beyond Profit

Wahala Thanthrige Navini Madhushani Perera[1], Jaki Erika[2]
[1] Doctoral School, Budapest University of Economics and Business, Hungary
[2] Budapest University of Economics and Business, Hungary

Abstract

This theoretical paper examines how intelligent knowledge management systems (IKMS) enable human–AI collaboration in financial decision-making and how such collaboration contributes to value creation beyond profit. The study addresses a theoretical gap at the intersection of behavioural finance, artificial intelligence in finance, knowledge management, and organisational innovation. Behavioural finance has demonstrated that financial judgement is shaped by biases, heuristics, emotions, and bounded rationality, while AI research in finance has largely emphasised automation, predictive accuracy, efficiency, and performance optimisation. However, less attention has been paid to how human cognition and algorithmic intelligence interact within organisational knowledge processes, and how this interaction reshapes financial decision-making in ways that affect trust, accountability, sustainability, and innovation. Using a conceptual framework-building approach, the paper positions IKMS as the central organisational layer through which human and AI cognition are integrated. Financial decision-making is conceptualised as a knowledge-intensive organisational process involving knowledge retrieval, knowledge integration, knowledge interpretation, and knowledge application. Within this process, AI contributes computational capacity, pattern recognition, prediction, and scalable knowledge synthesis, while human actors contribute contextual judgement, ethical reasoning, experience-based interpretation, and responsible oversight. The framework argues that human–AI collaboration becomes valuable not when AI replaces human judgement, but when IKMS enable complementarity, augmentation, validation, and continuous learning between human and machine intelligence. A central contribution of the paper is to show that AI does not eliminate behavioural bias in financial decision-making. Rather, biases may be redistributed, amplified, or transformed within hybrid human–AI knowledge systems. For example, AI-supported financial recommendations may reduce some forms of human overconfidence while creating new risks of automation bias, anchoring on algorithmic outputs, or excessive trust in model-generated classifications. The framework therefore identifies governance, explainability, trust calibration, ethical alignment, and feedback mechanisms as critical boundary conditions for responsible and effective human–AI collaboration. The paper conceptualises value creation beyond profit across four interrelated dimensions: cognitive value, reflected in improved decision quality, bias recalibration, and enhanced risk insight; institutional value, reflected in transparency, accountability, and stakeholder trust; societal value, reflected in ethical alignment, financial inclusion, sustainability, and long-term stability; and organisational innovation, reflected in improved knowledge recombination, adaptive learning routines, and more responsive financial decision processes. By integrating behavioural finance with intelligent knowledge management systems, the study contributes to a more comprehensive understanding of how AI-supported financial institutions can create trustworthy, sustainable, and innovation-oriented value beyond narrow economic performance.

Keywords: Intelligent knowledge management systems, human-AI collaboration, behavioural finance, organisational innovation, value creation beyond profit, sustainable and ethical AI
Paper type: Academic research paper (Theoretical)

1 Introduction

Over the past decade, artificial intelligence (AI) has significantly transformed financial decision-making (Aldasoro et al., 2025; Li et al., 2024), reshaping how information is processed, interpreted, and applied within financial institutions. In organisational terms, financial decision-making is not only an analytical activity but also a **knowledge-intensive process** (Hartvigsen, 1992) that depends on the acquisition, integration, interpretation, and use of dispersed forms of expertise and information (Richardson & Courtney, 2006). In line with Jarrahi et al. (2023), AI-enabled tools increasingly function as intelligent knowledge management systems by supporting core knowledge management activities such as retrieval, synthesis, sharing, augmentation, and application across organisational decision contexts (see also Böhm & Durst, 2025). This raises a broader question: how can human and AI cognition be integrated within organisational knowledge systems to generate not only performance gains but also innovation and sustainable value creation?

Despite the growing use of machine learning, predictive analytics, and algorithmic systems in finance, much of the literature continues to emphasise efficiency, predictive accuracy, and performance optimisation, often within a broader profit-oriented logic (Goodell et al., 2021; Pattnaik et al., 2024). This emphasis reflects the assumptions of traditional finance, where value is commonly framed in terms of market efficiency, shareholder

wealth, and risk–return optimisation (Fama, 1970; Stulz, 2025; Doeswijk et al., 2025). By contrast, behavioural finance (BF) has shown that financial decision-making is systematically shaped by cognitive limitations, heuristics, and emotional influences rather than strict rationality (Ackert, 2022; Kahneman & Tversky, 1979; Thaler, 1980). However, while BF has deepened understanding of human bias and judgment under uncertainty, it has largely treated decision-makers in isolation, with limited attention to how these behavioural dynamics are reconfigured when human cognition is embedded in AI-enabled organisational knowledge systems.

At the same time, research on AI in finance has often treated algorithms as tools of automation or optimisation (Goodell et al., 2021; Back, 2023), implicitly assuming that technological sophistication reduces or neutralises human bias (Hasan et al., 2024). Yet AI systems are not neutral actors: they may inherit, amplify, or reshape biases embedded in data, models, and institutional settings (Barocas & Selbst, 2016; Brynjolfsson & McAfee, 2017). As a result, an important theoretical gap remains. We still lack an integrated explanation of how human cognition and algorithmic intelligence interact within intelligent knowledge management systems, and how this interaction affects organisational knowledge processes, value creation, and innovation in financial institutions. This gap matters because contemporary financial organisations are increasingly expected to generate not only economic returns but also broader institutional and societal value, including transparency, trust, ethical alignment, and long-term sustainability (Freeman, 1984; Porter & Kramer, 2011; Dev et al., 2025; Jiang & Lin, 2026).

This paper addresses that gap by developing a theoretical framework that reconceptualises human–AI collaboration through a behavioural finance lens while placing intelligent knowledge management systems (IKMS) at the centre of the analysis. Rather than viewing AI as a substitute for human judgement, the paper conceptualises it as a cognitive complement that interacts with human cognition through an integrated human–AI cognition layer embedded in organisational knowledge processes. Within this framework, IKMS functions as the system layer through which human and AI capabilities are translated into core knowledge processes, namely knowledge retrieval, integration, interpretation, and application.

From this perspective, value does not arise solely from computational efficiency, but from the interaction between human interpretation, behavioural judgement, and AI-supported knowledge capabilities within hybrid knowledge systems. The paper argues that this interaction generates value beyond profit across three interrelated dimensions: i) cognitive value, reflected in improved decision quality, bias recalibration, and enhanced risk perception; ii) institutional value, reflected in transparency, accountability, and trust; and iii) societal value, reflected in ethical alignment, sustainability, and long-term financial stability. The framework further suggests that such systems may also support organizational innovation through improved knowledge integration, more adaptive decision processes, and enhanced learning routines in financial institutions.

Importantly, the paper also argues that AI does not eliminate behavioural bias; rather, it redistributes and reshapes bias within hybrid knowledge systems. This makes governance, explainability, trust, and ethical alignment central boundary conditions for effective and responsible human–AI collaboration. In this sense, the paper contributes not only to behavioural finance, but also to the literature on intelligent knowledge management systems by showing how human and AI cognition interact through organisational knowledge processes in financial decision-making. As a theoretical study, the paper does not test these relationships empirically; instead, it develops propositions to guide future research on human–AI collaboration, knowledge processes, value creation, and innovation in financial organisations.

2 Literature Review

2.1 Behavioural finance, AI in finance, and the limits of isolated decision models

Traditional financial theory assumes that rational agents operate in efficient markets where prices reflect available information (Fama, 1970; Lo, 2007). Recent work has refined this view by showing that market efficiency depends on information costs and distribution, rather than functioning as a frictionless benchmark (Lee, 2025). In this perspective, value creation is usually linked to profit maximisation, optimal allocation, and market performance. Behavioural finance (BF) challenged this assumption by showing that financial decisions are shaped by cognitive bias, bounded rationality, and emotions rather than strict rational calculation (Kahneman & Tversky, 1979; Thaler, 1980; Barberis & Thaler, 2003; Corzo, 2024). It identified recurrent patterns such as overconfidence, loss aversion, anchoring, and herd behaviour (Agudelo et al., 2024; Lonkani et al., 2025; Bai, 2024). However, BF has mainly examined human decision-makers, with less attention to how behavioural dynamics change when decisions are mediated by AI-enabled systems (Agudelo et al., 2024; Hasan et al., 2024; Klingbeil et al., 2024).

At the same time, AI-in-finance research has grown rapidly in areas such as algorithmic trading, credit scoring, fraud detection, and portfolio management (Goodell et al., 2021; Vuković et al., 2025). This literature largely emphasises efficiency, predictive accuracy, scalability, and personalization (Goodell et al., 2021; Pattnaik et al., 2024), often treating AI as an autonomous decision technology rather than part of a broader organisational knowledge environment. Emerging studies suggest that AI may inherit, amplify, or redistribute bias through data, models, and feedback loops (Barocas & Selbst, 2016; Aldasoro et al., 2025), while hybrid AI systems combine predictive power, contextual reasoning, and system design to improve financial decisions (Asemi et al., 2025). This reveals a limitation in both traditions: BF has paid limited attention to AI-mediated cognition, while AI-in-finance research has insufficiently examined how human judgement, bias, and organisational knowledge processes interact in hybrid decision systems.

2.2 Human–AI collaboration in organisational knowledge processes

Recent research has shifted from viewing AI as a substitute for human expertise toward understanding it as a complement to human cognition (Jarrahi, 2018; Simón et al., 2024; Cui et al., 2024; Byun, 2024; Bai, 2024; Back, 2023). This perspective emphasises complementarity, where human judgement and machine intelligence jointly improve decision-making outcomes (Daud et al., 2022). In organisational settings, this is best understood as human–AI collaboration embedded in knowledge processes. Humans provide contextual interpretation, ethical reasoning, practical judgement, and adaptive sensemaking, while AI contributes computational power, pattern recognition, retrieval speed, and consistency (Jarrahi, 2018; Cui et al., 2024). However, behavioural finance has paid limited attention to AI-mediated cognition, while AI-in-finance research has only recently examined how human judgement, bias, and organisational knowledge processes interact in hybrid decision systems (Hasan et al., 2024; Simón et al., 2024; Do Khac et al., 2025).

From this perspective, human–AI collaboration should be treated not merely as a technical arrangement but as a knowledge-organising mechanism for organisational value creation (Jarrahi, 2018; Simón et al., 2024). AI-enabled systems support knowledge augmentation through retrieval, synthesis, and pattern detection, while human actors remain central to interpretation, validation, and responsible application (Jarrahi et al., 2023; Cui et al., 2024). This is particularly relevant in finance, where decisions combine quantitative analysis with qualitative judgement under uncertainty (Back, 2023). Yet such interaction also creates behavioural and organisational challenges, including overreliance, algorithm aversion, shifting accountability, and trust calibration (Dietvorst et al., 2015; Do Khac et al., 2025). Thus, human–AI collaboration depends not only on better tools but on how cognitive and algorithmic capabilities are integrated within organisational knowledge systems.

2.3 Intelligent knowledge management systems, organisational learning and innovation

Intelligent knowledge management systems (IKMS) are central to understanding AI-supported decision-making in organisations. They can be understood as AI-enabled knowledge infrastructures that support knowledge retrieval, integration, interpretation, and application across organisational processes (Jarrahi et al., 2023; Böhm & Durst, 2025). In knowledge-intensive contexts such as financial institutions, IKMS do not merely automate information handling; they reshape knowledge flows, influence how expertise is combined, and structure decision support.

By improving access to dispersed knowledge, strengthening synthesis, and supporting interpretation, IKMS can enhance organisational learning and adaptive capability (Yan et al., 2026). Their relevance therefore extends beyond efficiency to organizational innovation, as they enable knowledge recombination, learning routines, responsive decision cycles, and better alignment between analytical systems and expert judgement (Simón et al., 2024; Cristache et al., 2025; An et al., 2024). Innovation should therefore be viewed as an organisational outcome of how knowledge is structured, shared, and applied through AI-enabled systems. However, these outcomes depend on system design and governance. Without transparency, human oversight, and knowledge validation, IKMS may reproduce weak judgements, reinforce flawed assumptions, or scale biased knowledge across the organisation (Olan et al., 2022; Rezaei, 2025).

2.4 Value creation beyond profit in AI-enabled knowledge systems

The concept of value creation in finance has traditionally been dominated by economic returns and shareholder wealth. More recent perspectives, however, emphasise that financial organisations are also expected to generate wider institutional and societal value, including trust, transparency, responsibility, and sustainability (Freeman, 1984; Porter & Kramer, 2011; Lim, 2024; Jiang & Lin, 2026; Pizzi et al., 2024). In AI-

enabled knowledge systems, value creation should therefore be understood as multidimensional rather than purely financial.

From a knowledge management perspective, value is generated not only through improved outcomes but through improved knowledge processes. When human judgement and AI capabilities are effectively integrated, organisations may create cognitive value through better decision quality and bias recalibration, institutional value through greater accountability and trust, and societal value through ethical alignment and long-term sustainability (Simón et al., 2024; Rezaei, 2025). In addition, such systems may also support organizational innovation by enabling more adaptive routines, stronger learning capability, and new forms of service, process, or business model development grounded in improved knowledge integration and application (Cristache et al., 2025; Yan et al., 2026; An et al., 2024). This suggests that value beyond profit in financial institutions is not separate from knowledge management; rather, it is increasingly produced through the design, governance, and application of intelligent knowledge systems.

2.5 Literature gap and theoretical positioning

Taken together, the literature points to four unresolved issues. First, behavioural finance explains human bias but does not adequately theorise AI as an active participant in organisational decision systems (Agudelo et al., 2024; Hasan et al., 2024). Second, AI-in-finance research explains technological capability but pays more limited attention to knowledge processes, human oversight, and behavioural dynamics (Goodell et al., 2021; Aldasoro et al., 2025; Rezaei, 2025). Third, the emerging literature on human–AI collaboration has not yet been sufficiently integrated with intelligent knowledge management systems and organisational learning (Simón et al., 2024; Do Khac et al., 2025; Yan et al., 2026). Fourth, the relationship between IKMS-enabled human–AI collaboration, value creation beyond profit, and organizational innovation remains undertheorized (Jarrahi et al., 2023; Simón et al., 2024; Cristache et al., 2025). This paper addresses these gaps by developing a framework that integrates behavioural finance with intelligent knowledge management systems in order to explain how human and AI cognition jointly shape knowledge-intensive financial decision-making and generate cognitive, institutional, and societal value beyond profit.

3 Conceptual Framework Development

3.1 Research design

This study adopts a conceptual design to develop an integrated framework explaining how intelligent knowledge management systems (IKMS) enable human–AI collaboration in financial decision-making and generate value beyond profit. This approach is appropriate because behavioural finance, AI-enabled decision-making, and knowledge management remain insufficiently integrated, especially regarding knowledge processes, governance conditions, and organizational innovation. Following guidance on conceptual article design, the study uses a framework-building approach to integrate fragmented streams (Jaakkola, 2020; Lim, 2026).

The paper applies a methodology suited to theory synthesis across behavioural finance, AI in finance, IKMS, knowledge processes, and multidimensional value creation (Jaakkola, 2020). It explains how human cognition and AI capabilities interact within financial institutions.

The framework was developed through three stages. First, literature synthesis identified constructs and tensions. Second, conceptual integration positioned human cognition and AI capabilities as interdependent sources of intelligence: humans contribute biases, heuristics, emotions, and expertise, while AI contributes data processing, prediction, learning, and reasoning. Their interaction links judgement and algorithmic intelligence in organisational decision-making (Jarrahi, 2018; Simón et al., 2024; Cui et al., 2024). Third, framework construction specified how these elements operate within IKMS. IKMS enables financial decision-making as an organisational knowledge process and supports knowledge-based value creation (Jarrahi et al., 2023; Simón et al., 2024). Human and AI cognition support retrieval, integration, interpretation, and application, enabling collaboration (Jarrahi et al., 2023; Rezaei, 2025).

3.2 Framework development

More specifically, the model distinguishes four interrelated knowledge processes through which human and AI capabilities can be translated into organisational decision support: i) knowledge retrieval refers to accessing internal and external data, information, and expertise relevant to financial decision-making, consistent with prior work showing that AI increasingly supports storage, retrieval, and situated knowledge application in

organisations (Jarrahi et al., 2023); ii) knowledge integration refers to the synthesis of heterogeneous inputs and the identification of patterns across sources, reflecting the broader challenge of combining dispersed organisational and domain knowledge within AI-enabled systems (Miedema et al., 2026); iii) knowledge interpretation refers to the sensemaking process through which integrated knowledge is contextualised, evaluated, and given practical meaning, particularly under conditions of uncertainty (Malik et al., 2025); iv) knowledge application refers to decision support, action recommendations, and the practical use of knowledge in financial decisions, in line with the KM literature that treats AI as increasingly relevant to knowledge use and implementation (Jarrahi et al., 2023; Rezaei, 2025).

Through these stages, IKMS does not merely automate information handling; it structures how knowledge is generated, combined, interpreted, and applied within financial institutions, while also introducing design and governance challenges that shape the quality of organisational outcomes (Jarrahi et al., 2023; Rezaei, 2025).

The framework further assumes that the operation of IKMS is shaped by a set of boundary conditions, namely governance, explainability, trust, and ethical alignment. These conditions influence whether human–AI cognition integration leads to effective and responsible collaboration or instead reproduces opacity, misalignment, or biased outcomes (Papagiannidis et al., 2025; Do Khac et al., 2025; Miedema et al., 2026). In this sense, the framework does not assume that AI eliminates behavioural bias. Rather, it recognises that biases may be reshaped, redistributed, or amplified within hybrid knowledge systems, making system design and oversight central to value creation (Barocas & Selbst, 2016; Xia et al., 2026; Kim et al., 2026).

Within this structure, human–AI collaboration is conceptualised as the immediate organisational mechanism emerging from the interaction of human cognition, AI capabilities, and IKMS-enabled knowledge processes. This view is consistent with prior work that treats human–AI collaboration not merely as technological interaction, but as an organisational arrangement through which human and algorithmic capabilities are combined for value creation (Jarrahi, 2018; Simón et al., 2024; Do Khac et al., 2025). This collaboration is characterised by complementarity, augmentation, and oversight, meaning that AI supports but does not replace human judgement. The model therefore treats collaboration not simply as technical coexistence, but as a knowledge-based mode of interaction in which human and AI capabilities jointly shape decision quality and organisational outcomes.

On this basis, the framework conceptualises value creation outcomes across four dimensions: i) cognitive value refers to improved decision quality, bias recalibration, and enhanced risk insight; ii) institutional value refers to transparency, accountability, and trust in AI-supported systems; iii) societal value refers to ethical alignment, sustainability, and long-term stability. In addition, the model includes a fourth dimension iv) organizational innovation as an important outcome, reflected in new processes, services, business models, and enhanced organisational learning and adaptive routines. This is consistent with emerging work suggesting that human–AI teaming and knowledge management can generate value not only through performance improvements, but also through trust-building, learning, and organisational adaptation (Simón et al., 2024; Cristache et al., 2025; Jiang & Lin, 2026). Accordingly, the framework assumes that intelligent knowledge management systems may generate not only better decisions, but also broader innovation-related benefits for financial institutions.

The framework further assumes that feedback and continuous improvement are integral to the operation of IKMS, such that the results of human–AI collaboration inform subsequent rounds of knowledge retrieval, integration, interpretation, and application (Yan et al., 2026; Zheng et al., 2025). In this sense, human–AI collaboration evolves through iterative learning, adaptive refinement, and ongoing recalibration of both system outputs and human judgement, consistent with organisational learning perspectives on AI capability development (Yan et al., 2026; Ritala et al., 2024).

Figure 1 presents this conceptual framework. It shows that value creation and organizational innovation do not emerge directly from either human cognition or AI alone, but from their interaction as embedded in intelligent knowledge management systems and mediated through organisational knowledge processes. The figure also highlights that feedback and continuous improvement are integral to the model, reinforcing the idea that human–AI collaboration evolves through ongoing learning, refinement, and adaptive innovation.

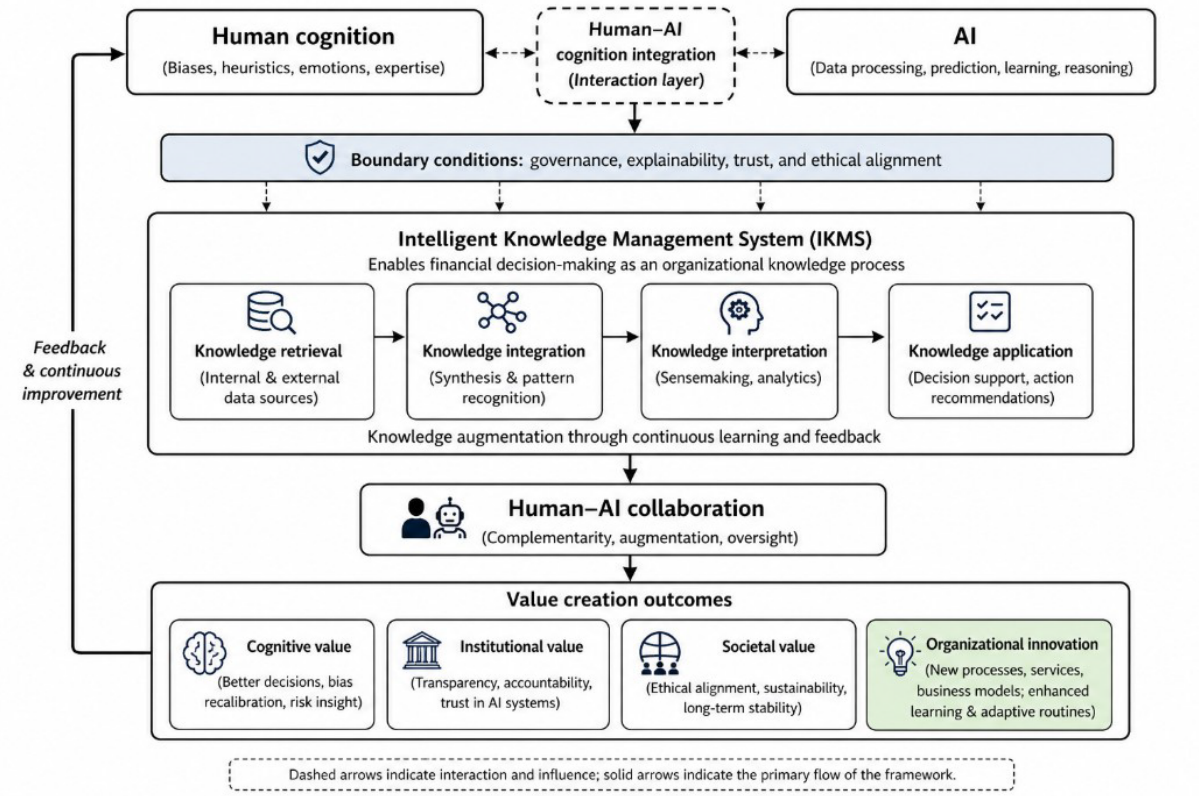


Figure 1: Conceptual framework

3.3 Proposition development

Based on the conceptual framework, the study develops five propositions specifying how human cognition, AI capabilities, intelligent knowledge management systems, organisational knowledge processes, and boundary conditions jointly shape value creation beyond profit in financial institutions.

P1: Human–AI cognitive complementarity within intelligent knowledge management systems enhances cognitive value through improved decision quality, enhanced risk insight, and bias recalibration.

P2: Intelligent knowledge management systems that effectively support knowledge retrieval, integration, interpretation, and application are more likely to enable effective human–AI collaboration in financial decision-making.

P3: Governance, explainability, trust, and ethical alignment act as boundary conditions that strengthen the positive relationship between human–AI collaboration and institutional and societal value.

P4: Feedback and continuous improvement mechanisms within intelligent knowledge management systems strengthen human–AI collaboration by enabling iterative learning and adaptive refinement of both system outputs and human judgement.

P**5:** Intelligent knowledge management systems that successfully integrate human and AI cognition are more likely to generate organizational innovation through improved knowledge recombination, adaptive learning routines, and more responsive decision processes.

These propositions suggest that the value of intelligent knowledge management systems lies not only in technological capability, but in how human and AI cognition are organised through knowledge processes, shaped by boundary conditions, and developed through iterative learning within financial institutions.

4 Discussion

The central contribution of this paper is to show that value creation in financial decision-making is not solely a technical or financial outcome, but the product of human–AI cognition integration embedded in intelligent knowledge management systems. By integrating behavioural finance, AI in finance, human–AI collaboration, and IKMS into a single framework, the study responds to a fragmented literature that has largely examined these domains separately. The framework suggests that financial decision-making should be understood as a knowledge-intensive organisational process in which value emerges through retrieval, integration, interpretation, and application rather than through technological capability alone. A first implication of the

framework is that human–AI collaboration is best understood as cognitive complementarity rather than substitution, as reflected in P1 and P2. AI contributes retrieval, synthesis, and pattern recognition, while human actors remain central to interpretation, validation, judgement, and responsible use. These propositions suggest that cognitive value depends both on the complementarity between human and AI capabilities and on the extent to which IKMS effectively support the underlying knowledge processes. Decision quality therefore becomes the outcome of hybrid cognition rather than isolated human judgement. A second implication is that bias must be analysed at the level of hybrid knowledge systems rather than at the level of individual decision-makers alone, which is captured in P3. The framework suggests that AI does not eliminate behavioural bias, but may redistribute, reshape, or amplify it through data, models, feedback loops, and patterns of organisational use. Governance, explainability, trust, and ethical alignment therefore function as boundary conditions that shape whether human–AI collaboration generates institutional and societal value or instead reinforces opacity and misalignment. A third implication is that human–AI collaboration should be treated as an iterative learning process rather than a static decision arrangement, as proposed in P4. By incorporating feedback and continuous improvement, the framework highlights that both system outputs and human judgement can be recalibrated over time. This feedback logic is especially important in financial institutions, where the quality of future decisions depends on whether prior outcomes are translated into organisational learning and system refinement. A fourth implication is that organizational innovation emerges as a systemic consequence of how IKMS structure knowledge processes, as formalised in P5. Better-designed human–AI knowledge systems do more than improve isolated decisions: they enable stronger knowledge recombination, more adaptive routines, and more responsive decision cycles. In this sense, organizational innovation is not external to value creation beyond profit, but one of its core organisational expressions.

These propositions generate three focused contributions. The paper extends behavioural finance toward hybrid knowledge systems, reframes AI in finance as an organisational knowledge process rather than a purely technical capability, and positions IKMS as the mechanism linking human–AI collaboration to multidimensional value creation and organizational innovation. For practice, the implication is equally clear: financial institutions should evaluate AI not only in terms of predictive performance, but also in terms of governance, knowledge quality, learning capacity, and the extent to which AI strengthens rather than weakens human judgement.

5 Conclusion

This paper examined how intelligent knowledge management systems (IKMS) enable human–AI collaboration in financial decision-making and how this collaboration contributes to value creation beyond profit. By integrating insights from behavioural finance, AI in finance, and knowledge management, the study showed that financial decision-making should be understood not only as a technical or analytical activity, but as a knowledge-intensive organisational process. The framework argues that AI functions not as a substitute for human judgement, but as a cognitive complement embedded in organisational knowledge processes of retrieval, integration, interpretation, and application.

The paper makes three main contributions. First, it extends behavioural finance by incorporating AI as an active participant in hybrid knowledge systems. Second, it reframes AI-supported decision-making in finance as a process shaped by knowledge governance, trust, explainability, and ethical alignment rather than by performance alone. Third, it shows that IKMS may generate not only cognitive, institutional, and societal value, but also organizational innovation through improved learning, knowledge recombination, and adaptive decision routines. Overall, the study suggests that the future of financial decision-making lies not in full automation, but in the development of well-governed intelligent knowledge management systems that support more trustworthy, adaptive, and innovation-oriented financial organisations.

6 Ethics Declaration

Ethical clearance was not required for this paper.

7 AI Declaration

This paper used Grammarly software to check the English language, as English is not the authors' native language.

References

Ackert, L. F. (2022). Behavioral finance: Psychology, decision-making, and markets. *Cengage.*

Agudelo Aguirre, R. A., & Agudelo Aguirre, A. A. (2024). Behavioural finance: Evolution from the classical theory and remarks. *Journal of Economic Surveys*, *38*(2), 452–475.

Aldasoro, I., Gambacorta, L., Korinek, A., Shreeti, V., & Stein, M. (2025). *Intelligent financial system: How AI is transforming finance*. *Journal of Financial Stability*, *81*, 101472.

An, M., Lin, J., & Luo, X. R. (2024). The impact of human AI skills on organizational innovation: The moderating role of digital organizational culture. *Journal of Business Research*, *182*, 114786.

Asemi A, Sebrek SS, Pérez Garrido B (2025;), "Transforming financial decision-making with hybrid artificial intelligence (AI)". *Management Decision*, Vol. ahead-of-print No. ahead-of-p rint.

Back, C., Morana, S., & Spann, M. (2023). When do robo-advisors make us better investors? The impact of social design elements on investor behavior. *Journal of Behavioral and Experimental Economics*, *103*, 101984.

Bai, Z. (2024). Examining the association between robo-advisory and perceived financial satisfaction. *Review of Behavioural Finance*, *16*(4), 668–681.

Barberis, N., & Thaler, R. (2003). A survey of behavioral finance. In G. M. Constantinides, M. Harris, & R. M. Stulz (Eds.), *Handbook of the Economics of Finance* (Vol. 1, pp. 1053–1128). Elsevier.

Barocas, S., & Selbst, A. D. (2016). Big data's disparate impact. *California Law Review*, 104(3), 671–732.

Böhm, K., & Durst, S. (2025). *Knowledge management in the age of generative artificial intelligence: From SECI to GRAI*. *VINE Journal of Information and Knowledge Management Systems*, *56*(1), 106–121.

Brynjolfsson, E., & McAfee, A. (2017). The business of artificial intelligence. *Harvard Business Review.*

Byun, S.-J., Cho, S., & Kim, D.-H. (2024). Can a machine learn from behavioral biases? Evidence from stock return predictability of deep learning models. *Journal of Behavioral and Experimental Finance*, *41*, 100881.

Corzo, T. (2024). *Behavioral finance in a hundred keywords*. *Heliyon*, *10*(23), e40603.

Cristache, N., Croitoru, G., & Florea, N. V. (2025). The influence of knowledge management on innovation and organizational performance. *Journal of Innovation & Knowledge*, *10*(5), 100793.

Cui, H., Hertwig, R., & Rahwan, I. (2024). *AI-enhanced collective intelligence*. *Trends in Cognitive Sciences*, *28*(12), 1048–1060.

Daud, S. N. M., Khalid, A., & Azman-Saini, W. N. W. (2022). FinTech and financial stability: Threat or opportunity?. *Finance Research Letters*, *47*, 102667.

Dev, D., Sharma, G. D., Gupta, M., & Tiwari, A. K. (2025). *Sustainable finance in action: A comprehensive framework for policy and practice integration*. *International Review of Economics & Finance*, *103*, 104511.

Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology*: General, 144(1), 114–126.

Do Khac, L. T., Malik, A., Nguyen, B., & Papagiannidis, S. (2025). *Towards an integrative model of organizational human-AI collaboration: A trust-centered perspective*. *Technology in Society*, *83*, 103028.

Doeswijk, R. Q., Lam, T., & Swinkels, L. A. P. (2025). *The risk and reward of investing*. *Journal of Banking & Finance*, *175*, 107343.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *Journal of Finance*, 25(2), 383–417.

Freeman, R. E. (1984). Strategic management: A stakeholder approach. *Pitman*.

Goodell, J. W., Kumar, S., Lim, W. M., & Pattnaik, D. (2021). Artificial intelligence and machine learning in finance: Identifying foundations, themes, and research clusters from bibliometric analysis. *Journal of Behavioral and Experimental Finance*, *32*, 100577.

Hartvigsen, G. (1992). *Limitations of knowledge-based systems for financial analysis*. *Expert Systems with Applications*, *4*(3), 255–268.

Hasan, Z., Kamil, N. M., & Mustafa, A. (2024). *Can artificial intelligence (AI) manage behavioural biases in the financial decision-making process? A systematic literature review*. *The North American Journal of Economics and Finance*, *70*, 102073.

Jaakkola, E. (2020). *Designing conceptual articles: Four approaches*. *AMS Review*, *10*(1–2), 18–26.

Jarrahi, M. H. (2018). Artificial intelligence and the future of work: Human–AI symbiosis in organizational decision making. Business Horizons, 61(4), 577–586.

Jarrahi, M. H., Askay, D., Eshraghi, A., & Smith, P. (2023). Artificial intelligence and knowledge management. Business Horizons, 66(1), 87–99. https://doi.org/10.1016/j.bushor.2022.03.002

Jiang, Y., & Lin, H. (2026). Bridging sustainability and value co-creation: The financial and strategic role of ESG performance. International Review of Economics & Finance, 105, 104861.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2), 263–291.

Kim, W., & Ryoo, Y. (2026). The literacy paradox: How AI literacy amplifies biases in evaluating AI-generated news articles. *International Journal of Information Management*, *86*, 102992.

Klingbeil, A., Syska, A., Rauhut, H., & Heekeren, H. R. (2024). *Trust and reliance on AI—An experimental study*. *Computers in Human Behavior*, *160*, 108240.

Lee, C. M. C. (2025). *Market efficiency with costly information*. SSRN Electronic Journal.

Li, Y., Zhong, H., & Tong, Q. (2024). *Artificial intelligence, dynamic capabilities, and corporate financial asset allocation*. *International Review of Financial Analysis*, *96*, 103773.

Lim, T. (2024). Environmental, social, and governance (ESG) and artificial intelligence in finance: State-of-the-art and research takeaways. *Artificial Intelligence Review*, *57*(4).

Lim, W. M. (2026). *Theory and theory development: Guidelines for establishing theoretical rigor in business research*. *Journal of Business Research*, *202*, Article details as indexed online.

Lo, A. W. (2007). *Efficient markets hypothesis*. In S. N. Durlauf & L. E. Blume (Eds.), *The New Palgrave Dictionary of Economics* (2nd ed.). Palgrave Macmillan.

Lonkani, R., Changchit, C., Cutshall, R., Treerotchananon, A., & Prasertsoontorn, T. (2025). The effect of investors' emotional and depressive states on perceived returns and risk. *Journal of Asian Business and Economic Studies*, *32*(3), 134-146.

Malik, M., Andargoli, A., Tallon, P., & Wickramasinghe, N. (2025). An organizational sensemaking theorizing of how firms construct digitally enabled strategic agility. *Information & Management*, *62*(4), 104130.

Miedema, E., Waschull, S., & Emmanouilidis, C. (2026). Towards trustworthy artificial intelligence for decision-making: A lifecycle perspective on knowledge-and data-driven artificial intelligence systems. *Computers in Industry*, *174*, 104409.

Olan, F., Arakpogun, E. O., Suklan, J., Nakpodia, F., Damij, N., & Jayawickrama, U. (2022). Artificial intelligence and knowledge sharing: Contributing factors to organizational performance. *Journal of Business Research*, *145*, 605-615.

Papagiannidis, E., Mikalef, P., & Conboy, K. (2025). Responsible artificial intelligence governance: A review and research framework. *The Journal of Strategic Information Systems*, *34*(2), 101885.

Pattnaik, D., Kumar, S., Burton, B., Lim, W. M., & Dwivedi, Y. K. (2024). *Applications of artificial intelligence and machine learning in the banking, financial services, and insurance sector: A bibliometric review*. *Heliyon*, *10*(1), e24004.

Pizzi, S., Venturelli, A., & Caputo, F. (2024). *Restoring trust in sustainability reporting: The enabling role of the external assurance*. *Current Opinion in Environmental Sustainability*, *70*, 101437.

Porter, M. E., & Kramer, M. R. (2011). Creating shared value. *Harvard Business Review*, 89(1–2), 62–77.

Rezaei, M. (2025). Artificial intelligence in knowledge management: Identifying and addressing the key implementation challenges. *Technological Forecasting and Social Change*, *217*, 124183.

Richardson, S. M., & Courtney, J. F. (2004/2005 online; journal record 2006). *Theoretical principles for knowledge management system design*. Information Systems Frontiers, 8(2), 147–157.

Simón, C., Revilla, E., & Sáenz, M. J. (2024). Integrating AI in organizations for value creation through Human-AI teaming: A dynamic-capabilities approach. *Journal of Business Research*, *182*, 114783.

Stulz, R. M. (2025). *The lessons of Michael C. Jensen*. *Journal of Corporate Finance*, *94*, 102846.

Thaler, R. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior & Organization*, 1(1), 39–60.

Vuković, D. B., Popp, J., Fragkou, P., Cumming, D., Molnár, P., & Khan, K. A. (2025). *AI integration in financial services: A systematic review of trends, applications, and implications*. *Humanities and Social Sciences Communications*, *12*, Article 386.

Xia, H., Chen, H., Zhang, J. Z., & Kamal, M. M. (2026). Exploring the impact of responsible AI governance on corporate performance: A quasi-natural experiment. *Technological Forecasting and Social Change*, *223*, 124425.

Yan, J., Leidner, D., & Benbya, H. (2026). *Organizational learning with artificial intelligence*. *Long Range Planning*, *59*(1), 103066.

Zheng, J., Richter, A., Hong, Y., & Holmström, J. (2025). The agentic shift: Making human-AI coordination work by addressing two critical junctures. *Business Horizons*.

Assessing Knowledge Management Maturity in Public Sector Organizations: Model Development and Case-Based Validation

Ingrida Girnienė, Laura Valiukevičienė
Vilnius University, Faculty of Communication, Vilnius, Lithuania

Abstract

Knowledge management (KM) maturity has become increasingly relevant as organizations seek to improve operational reliability, support digital transformation, and prepare for the responsible use of artificial intelligence (AI). However, existing KM maturity models are often discussed as generic frameworks and are less frequently adapted and empirically applied in specific organizational and public-sector contexts. This paper develops and applies a context-adapted KM maturity assessment model in a Lithuanian public-sector transport organization. The study addresses the question of how KM maturity can be purposefully assessed through a model that is sensitive to organizational, sectoral, and contemporary digital-transformation conditions, and what level of maturity characterizes the selected case organization. The model was constructed through a comparative analysis of established KM maturity models, including Siemens KMMM, G-KMMM, APQC, KNM 2.0, and IT-CMF, and was aligned with ISO 30401:2018 principles. In response to the growing role of AI in KM, the model distinguishes between innovation-enabled high maturity and AI-supported maturity at the highest stage. The empirical research applies a case-study strategy combining strategic document analysis and semi-structured interviews with managers from different functional areas. Findings show that KM is present in the organization, but is not explicitly formalized as an independent strategic priority. The strongest dimensions are culture, people, and technology: employees generally experience a cooperative environment, learning systems are available, and multiple digital tools support knowledge storage and communication. However, KM processes remain fragmented, knowledge ownership is unclear, search and taxonomy are weak, and KM performance indicators are absent. The paper contributes by proposing a diagnostic model that is both descriptive and prescriptive: it assesses the current maturity state and indicates development priorities for progression. Managerially, the study highlights the need to formalize KM policy, clarify ownership, integrate KM objectives with technology and AI initiatives, and introduce KM performance measurement.

Keywords: knowledge management maturity; public sector; maturity model; digital transformation; artificial intelligence

Paper type: Academic Research Paper

1 Introduction

Knowledge has long been recognized as a strategic organizational resource, but its managerial importance has intensified amid digital transformation and artificial intelligence. Organizations increasingly rely on accurate, accessible, and reusable knowledge to make decisions, redesign processes, support service quality, and retain expertise when employees move between roles or leave. In the public sector, where continuity, accountability, and service reliability are especially important, knowledge management (KM) is not only a managerial tool, but also a condition for institutional resilience and public value creation.

The relevance of KM maturity has grown because digital and AI-enabled transformation depends on the availability, quality, and governance of organizational knowledge. Recent discussions of AI adoption emphasize that value is achieved not only through technology acquisition, but through redesigned workflows, data and knowledge governance, human validation, measurable outcomes, and organizational adoption practices (McKinsey & Company, 2025; Gartner, 2024). In this sense, KM maturity is not treated as a separate technological capability in this paper. It is conceptualized as an organizational capability that conditions how well technology and AI can be used for knowledge creation, storage, sharing, and application.

Although KM maturity models offer structured ways to assess such capability, the literature also identifies several limitations. Many models have been designed for broad or large-scale corporate use, emphasize only selected dimensions, and are difficult to apply directly in specific national, sectoral, or organizational contexts (Escrivão and Da Silva, 2019; Bougoulia and Glykas, 2023). Empirical applications in public-sector organizations remain comparatively limited. In Lithuania and the Baltic context, KM maturity assessment is particularly underdeveloped, with only isolated prior applications and limited evidence about how maturity models should be adapted to local organizational realities (Šajeva and Jucevičius, 2010).

This study therefore aims to develop and empirically validate a context-adapted KM maturity assessment model that enables systematic evaluation of KM practices in a public-sector organization and supports the identification of development priorities. The study is situated in organizational, public-sector, and contemporary

digital-transformation contexts. It is an organizational context, because the model is adapted to the selected case organization and its existing practices. It is a public-sector context, because the organization performs strategically important passenger transport functions and operates under public accountability and service-continuity expectations. It is a contemporary digital-transformation context because the model explicitly considers technology, performance measurement, and AI-supported maturity.

The main research question in the paper is: how can KM maturity be assessed in an organization through a context-appropriate maturity model? The contribution is both theoretical and practical. Theoretically, the paper integrates classical KM maturity logic with ISO 30401:2018 and digital-transformation considerations. Practically, it proposes a diagnostic model that identifies strengths, weaknesses, and improvement pathways.

2 Literature Background

KM is commonly understood as a set of organizational processes through which knowledge is created, stored, shared and applied to support organizational objectives (North and Kumta, 2018; Alavi and Leidner, 2001; Dalkir, 2021). Early KM research emphasized the socio-technical nature of KM systems: technologies can support knowledge processes, but their effectiveness depends on people, culture, organizational structure, and managerial priorities. Nonaka and Takeuchi (1995) similarly argued that knowledge creation requires interaction between tacit and explicit knowledge, making organizational culture and social practices central to KM performance.

Maturity models emerged as tools to assess how systematically an organization manages a particular capability. The Capability Maturity Model logic influenced many KM maturity frameworks by introducing staged development from informal and reactive practices to optimized and continuously improved processes (Paulk et al., 1993; Rios, Rodríguez-Elias and Pino, 2011). KM maturity models apply this logic to knowledge-related activities, helping organizations diagnose their current state and plan improvement.

Several established KM maturity models were analyzed in this study. The Siemens KMMM by Ehms and Langen (2002) offers a socio-technical assessment across multiple areas, including strategy, people, culture, knowledge structures, technology, and processes. The G-KMMM integrates people, processes, and technology and defines sequential maturity levels from initial to optimizing (Pee and Kankanhalli, 2009). The APQC model emphasizes standardized assessment and benchmarking across areas such as people, process, content, technology, and strategy (APQC, 2016; Hubert and Lemons, 2019). KNM and KNM 2.0 offer a more detailed assessment with dimensions and indicators that can be converted into maturity indices (Hsieh, Lin and Lin, 2009; Hsieh et al., 2020). IT-CMF extends the discussion toward the digital age by emphasizing that KM maturity must account for analytics, digital tools, information overload and the capacity to learn dynamically (Thornley et al., 2016).

The comparison of these models indicates some important outcomes:

- people, processes and technology are foundational and appear in nearly all models;
- culture, strategy and performance measurement are often decisive for the sustainability of KM, but are not always sufficiently explicit;
- digital transformation and AI require maturity models to move beyond technology availability toward integration, governance, quality control, human oversight and measurable value.

ISO 30401:2018 reinforces this broader view by framing KM as a management system aligned with organizational context, leadership, planning, support, operation, evaluation and improvement (ISO, 2018; Carlucci et al., 2022).

The literature also shows a gap between model design and empirical use. Escrivão and Da Silva (2019) identify fragmentation and the need for models that reflect organizational complexity rather than isolated technological indicators. Bougoulia and Glykas (2023) argue for more holistic frameworks aligned with ISO 30401 and for stronger implementation orientation. This is especially relevant for public-sector organizations, where KM may be embedded in HR, quality, digitalization or operational routines without being labelled as KM.

3 Model Development

The adapted model developed in this study includes seven dimensions: people, processes, technology, culture, strategy and management, performance management, and results (Figure1). The model is context-adapted because these dimensions were selected through a synthesis of classical KM maturity models, ISO 30401:2018 requirements and the empirical characteristics of the selected organization. It does not assume that KM must be managed through a separate department; instead, it assesses whether KM practices are sufficiently visible, coordinated, measurable and value-generating across organizational routines.

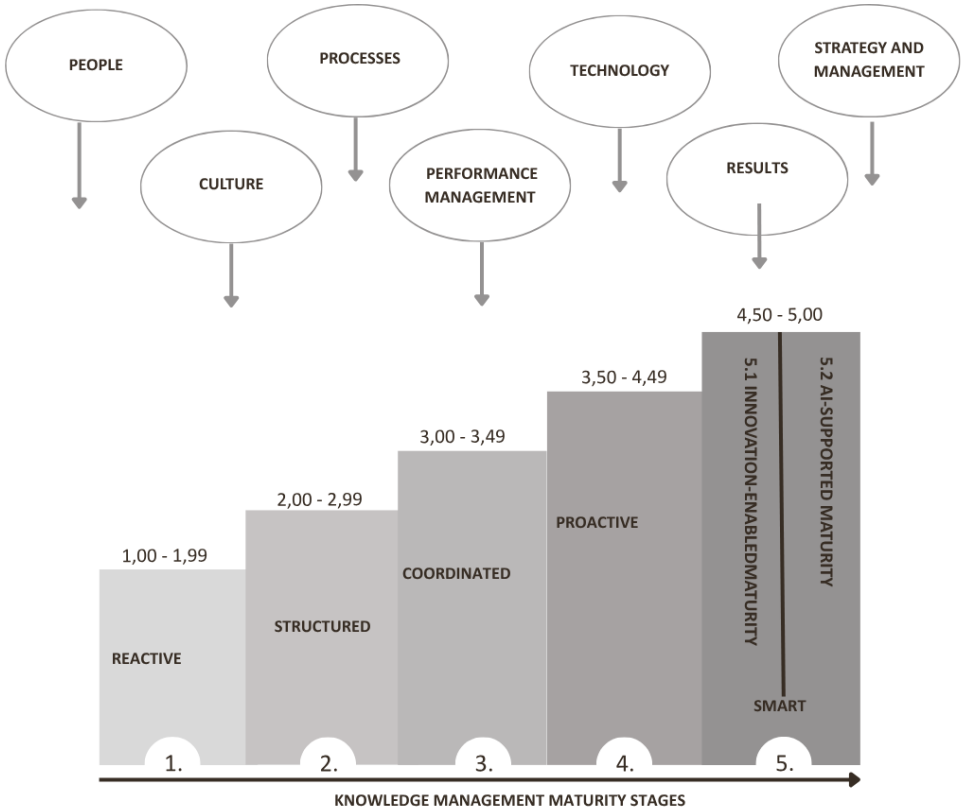


Figure 1 : Knowledge Management adapted model

The people dimension assesses competence development, motivation, roles, leadership, and knowledge-sharing behaviour. The processes dimension assesses the documentation, storage, searchability, repeatability, updating, and control of knowledge processes. The technology dimension evaluates the availability, usability, and integration of digital tools, including AI-related initiatives. The culture dimension assesses trust, openness, collaboration, learning from mistakes, and managerial role modeling. The strategy and management dimension evaluates the alignment of KM with organizational objectives, policy, governance, and resources. The performance management dimension assesses whether KM indicators, audits, and feedback mechanisms exist and are used for improvement. The results dimension evaluates whether KM creates visible organizational value, for example, faster onboarding, improved decisions, fewer errors, process efficiency, or innovation.

The maturity scale consists of five main stages. Level 1, reactive, describes organizations where KM is largely unrecognized, and knowledge is managed informally. Level 2, structured, indicates the existence of initial practices, documents or tools, but with limited coordination. Level 3, coordinated, describes organizations where core KM practices exist across several areas but are not fully integrated strategically or measured consistently. Level 4, proactive, indicates systematic KM practices, clear responsibilities, managerial support and continuous improvement. Level 5 describes smart maturity and is divided conceptually into two subtypes: 5.1 innovation-enabled maturity, where KM is integrated into innovation and advanced digital tools; and 5.2 AI-supported maturity, where AI is embedded in knowledge search, classification, reuse, recommendation, validation or real-time application. The sublevels are not scored separately in this paper; they clarify the difference between classical high maturity and AI-enabled KM maturity.

The model is primarily diagnostic, but also prescriptive. It is descriptive because it assesses the current level of maturity in each dimension. It is prescriptive because each level implies a development path: for example, an organization at Level 2 in performance management must define KM indicators and evaluation routines before it can claim Level 3 or Level 4 maturity.

4 Methodology

The empirical study adopts a qualitative case-study strategy, supported by structured scoring. A case study is appropriate because KM maturity is a complex organizational phenomenon embedded in routines, practices, documents, and managerial interpretations (Yin, 2018). The selected case is a Lithuanian public-sector

organization. The case was selected purposively for its strategic public-service role, operational complexity, digitalization agenda, emphasis on service quality, and reliance on coordinated knowledge across functions.

Two data sources were used (Figure 2). First, a strategic document analysis was conducted using the 2024 annual report, the 2025 first-half report, and the 2029 organizational strategy. The document analysis identified explicit and implicit references to the seven model dimensions. Repeated content across at least two documents was treated as stronger evidence of stable organizational practice, while isolated references were considered supplementary. Second, seven semi-structured interviews were conducted with managers from different functional areas: project management, operational efficiency, customer service, business support, international development, technical project development, and human resources. The informants represented both strategic and tactical perspectives and had organizational experience ranging from less than one year to eighteen years.

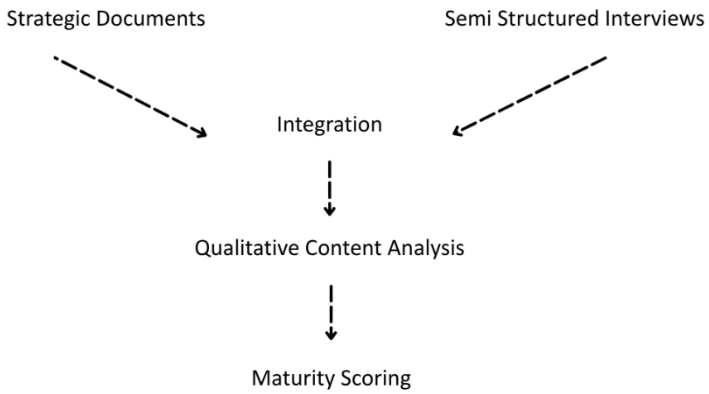


Figure 2: Data source integration

Interview questions were structured around the seven model dimensions. Seven interviews were conducted remotely and in person, recorded, and transcribed. The analysis followed qualitative content analysis. Evidence was coded into categories and subcategories, such as knowledge storage, role clarity, learning systems, searchability, knowledge-sharing culture, strategic alignment, and performance measurement. The qualitative findings were then translated into maturity scores for each dimension using the adapted maturity scale. The final score was calculated as the arithmetic mean of the dimensional scores. This scoring approach was used not to claim statistical generalizability, but to make the diagnostic interpretation transparent and comparable across dimensions.

5 Findings

The strategic document analysis shows that KM is not explicitly articulated as an independent strategic priority in the organization. There is no separate KM policy, no explicit KM objectives, and no formalized KM performance measurement system. Nevertheless, several KM-related elements appear implicitly in the organization's strategic language and operational initiatives. The strongest documentary evidence relates to people, culture, and technology. The documents repeatedly emphasize improvement, cooperation, employee development, inclusion, and digitalization. Programs such as onboarding, recognition initiatives, learning platforms, and competence evaluation indicate that knowledge-related activities are embedded in HR and organizational development routines.

The people dimension received a score of 3, corresponding to a coordinated level. The organization has a competence management center, learning systems, mandatory and additional training, mentoring-like practices, and recognition mechanisms. Managers confirm the existence of a learning management system and formal competence-development practices. However, knowledge ownership is not always clear, and some employees, particularly in operational roles, may hesitate to document their knowledge because tacit expertise is perceived as a means of protecting individual replaceability. This creates a risk of knowledge loss and shows that people-related KM practices, while present, are not yet fully systematic.

The processes dimension received a score of 2.5, indicating a maturity between structured and coordinated. Formal processes exist for onboarding, training, and repeated tasks, and employees use templates, forms, and established documents. At the same time, knowledge storage and retrieval remain fragmented. Knowledge is stored across SharePoint, Teams, document management systems, and other tools, but there is no unified taxonomy or reliable search logic. Managers report that keyword searches do not always lead to the most current or relevant information. Knowledge capture is being addressed through planned pilot initiatives, such as

a professional-mastery project, but these initiatives are not yet institutionalized. Process control is weak: KM processes are not evaluated as KM processes, and general audits do not substitute for systematic knowledge-process assessment.

The technology dimension received a score of 3, indicating a coordinated level of maturity. The organization uses multiple digital tools, including a document management system, learning management system, SharePoint, Microsoft Teams, SAP-related competence tools, and JIRA. These tools enable information storage, communication, learning, and task coordination. Technology, therefore, supports KM, but mainly as infrastructure. The tools are not sufficiently integrated, and users experience search, usability, and taxonomy problems. AI use is emerging but not yet embedded in KM. Some managers use Microsoft Copilot individually, and a project for transforming the knowledge base with AI is planned. However, AI currently supports individual productivity rather than an integrated organizational KM system.

The culture dimension received the highest score, 4, corresponding to proactive maturity. Interviews indicate a generally open, friendly, and cooperative environment. Managers describe employees as willing to help one another, and leadership is perceived as modeling information sharing. Recognition practices and public acknowledgment of knowledgeable employees support positive knowledge-sharing norms. This finding aligns with strategic documents emphasizing cooperation, development, and respectful working conditions. However, the cultural score is not 5 because knowledge hoarding persists in some roles and because the link between cultural values and formal KM routines is incomplete.

The strategy and management dimension received a score of 2, corresponding to structured maturity. The organization has a long-term strategy, clearly communicated values, and strategic development cycles, but KM is not named as a strategic priority. Interviewees explicitly note that there is no separate KM policy. Knowledge sharing is visible through values and inter-organizational cooperation, including learning from partner organizations, but these practices remain embedded in general strategy rather than governed as KM. Resources are not allocated specifically to KM, and ownership of some knowledge processes remains unclear.

The performance management dimension received the lowest score, 1, corresponding to reactive maturity. The organization uses feedback tools and competence-evaluation practices and collects general operational indicators. However, there are no KM-specific indicators, no systematic evaluation of knowledge creation, sharing, reuse, or loss, and no measurement of the value created by KM initiatives. This creates a major barrier to maturity progression. Without measurement, the organization cannot reliably determine whether knowledge processes improve decision-making, reduce repeated errors, accelerate onboarding, or support innovation.

The results dimension received a score of 2, indicating a structured level of maturity. Managers recognize knowledge and experience as highly valuable and directly related to strategic goals. One interviewee emphasized that achieving the organization's strategic aspiration requires improving knowledge transfer and KM practices. However, KM results are not measured directly. Benefits are identified only indirectly through general performance, service improvement, competence development, and strategic projects. Knowledge risks are also visible: outdated content, weak searchability, and insufficient documentation may reduce the reliability and speed of decision-making.

Overall, the organization's KM maturity is assessed at 2.5 out of 5, indicating a structured and coordinated level. The organization has important foundations for KM: learning systems, digital tools, cooperative culture, and managerial awareness. Yet these foundations remain only partly integrated. KM is not formalized strategically, processes are fragmented, technology is not fully interoperable, and performance measurement is absent.

Table 1: Summary of maturity scores

Dimension	Score	Maturity interpretation	Main evidence
People	3.0	Coordinated	Competence management, LMS, training, and mentoring are in place; role clarity and knowledge capture remain uneven.
Processes	2.5	Structured/coordinated	Templates, onboarding, and training processes exist; storage, search, taxonomy, and control are fragmented.
Technology	3.0	Coordinated	Multiple tools support KM; integration, usability, and AI embedding remain limited.
Culture	4.0	Proactive	Cooperative, open environment and leadership support; some knowledge hoarding persists.
Strategy and management	2.0	Structured	A long-term strategy and values exist, but KM policy, objectives, and resources are not explicitly defined.

Dimension	Score	Maturity interpretation	Main evidence
Performance management	1.0	Reactive	General indicators and feedback exist, but KM indicators and evaluation are absent.
Results	2.0	Structured	Knowledge value is recognized, but KM benefits and risks are not systematically measured.

6 Discussion

The findings confirm that KM maturity in the selected public-sector organization is not absent, but it is implicit, fragmented, and unevenly developed. This supports the argument that KM in contemporary organizations often appears through HR, quality management, digitalization, and operational routines rather than through explicit KM policies. The case, therefore, illustrates why generic maturity models need contextual adaptation. A model limited to technology, for example, would overestimate maturity because many tools are available. A model limited to culture would also overestimate maturity because the organization demonstrates strong collaboration. Conversely, a model incorporating strategy, performance management, and results reveals that KM has not yet become a systematically governed capability.

The study also clarifies the relationship between KM maturity, AI, and digital transformation. AI is represented in the model in two ways. First, it appears in the technology dimension, where the empirical assessment examines whether AI tools are used, integrated, and connected to organizational knowledge repositories. Second, AI is incorporated into the highest maturity level, where AI-supported KM is distinguished from general innovation-enabled maturity. In the case of an organization, AI is not yet a maturity driver; it is an emerging initiative. This finding is important because it avoids treating AI adoption as proof of maturity. AI-supported KM requires reliable knowledge bases, governance, taxonomy, quality control, human validation, and performance measurement. Without these foundations, AI may accelerate access to poorly governed knowledge rather than improve KM maturity.

The novelty of the paper lies in three elements. First, it proposes a seven-dimensional maturity model that combines classical KM maturity dimensions with ISO-aligned governance, performance, and results dimensions. Second, it adapts the maturity scale to digital transformation by separating innovation-enabled and AI-supported maturity at the highest level. Third, it empirically applies the model in a Lithuanian public-sector context, where KM maturity studies are limited. This responds to calls for more grounded and context-sensitive KM maturity research (Escrivão and Da Silva, 2019; Bougoulia and Glykas, 2023).

Theoretical implications arise from the model's multidimensional character. The study demonstrates that KM maturity should be evaluated as an organizational configuration rather than a single capability. High maturity requires alignment among culture, process, technology, strategy, and measurement. The case shows that strong culture and technology do not compensate for weak performance management. This finding supports socio-technical views of KM and extends them by emphasizing measurement and AI-readiness as contemporary maturity conditions.

Managerial implications are also clear. The organization should formalize KM as a strategic objective or policy area, without necessarily creating a separate bureaucratic unit. Knowledge ownership should be clarified for critical documents, processes, and repositories. A unified taxonomy and semantic search approach should be developed to improve knowledge accessibility. AI initiatives should be aligned with KM objectives and governed by quality, responsibility, and human-validation rules. KM indicators should be introduced, such as time required to find critical knowledge, onboarding duration, reuse of documented lessons, frequency of content updates, knowledge-loss incidents, and employee perceptions of knowledge accessibility. These steps would move the organization from a structured/coordinated maturity level to a proactive maturity.

7 Conclusions

This paper developed and applied a context-adapted KM maturity assessment model in a public-sector organization. The study shows that KM maturity can be purposefully assessed by combining established maturity-model logic with context-sensitive dimensions and empirical evidence from documents and managerial interviews. The proposed model is both descriptive and prescriptive: it identifies the current maturity level and indicates what must be improved to progress to higher maturity.

The organization has a strong collaborative culture, established learning practices, and multiple digital tools. However, KM is not explicitly defined in strategy, processes are fragmented, knowledge search and taxonomy are weak, AI is not yet embedded in KM, and KM performance is not measured. The most important maturity barrier is therefore not the absence of knowledge-related activity, but the absence of systematic integration and evaluation.

The study contributes to KM maturity literature by offering a model suitable for public-sector and digitally transforming organizations. AI-supported KM requires prior development of governance, knowledge quality, process integration, and measurement. Future research could apply the model across multiple organizations, compare public and private sector cases, and test the scoring instrument quantitatively using a larger sample.

8 Ethics Declaration

The empirical research was conducted with voluntary participation from managers. Interviewees were anonymized, and findings are reported in aggregated form to avoid disclosure of personal or sensitive organizational information.

9 AI Declaration

AI tools were used to support language editing, structuring, and academic drafting based on the author-provided research material. The author is responsible for verifying the accuracy, validity, and appropriateness of the final paper.

References

Alavi, M. and Leidner, D.E. (2001) 'Review: Knowledge management and knowledge management systems: Conceptual foundations and research issues', MIS Quarterly, 25(1), pp. 107-136.

APQC (2016) KM Capability Assessment Tool: Orientation Session. Houston: American Productivity and Quality Center.

Bougoulia, E. and Glykas, M. (2023) 'Knowledge management maturity assessment frameworks: A proposed holistic approach', Knowledge and Process Management, 30(4), pp. 355-386.

Carlucci, D., Kudryavtsev, D., Santarsiero, F., Lagrutta, R. and Garavelli, A.C. (2022) 'The ISO 30401 Knowledge Management Systems: A new frame for managing knowledge. Conceptualisation and practice', Knowledge Management Research & Practice, 20(6), pp. 975-986.

Dalkir, K. (2021) Knowledge Management in Theory and Practice. 4th edn. Cambridge, MA: MIT Press.

Ehms, K. and Langen, M. (2002) Holistic Development of Knowledge Management with KMMM. Munich: Siemens AG.

Escrivão, G. and Da Silva, S.L. (2019) 'Knowledge management maturity models: Identification of gaps and improvement proposal', Gestão & Produção, 26(3).

Gartner (2024) What is AI-ready data?

Hsieh, P.J., Lin, B. and Lin, C. (2009) 'The construction and application of knowledge navigator model: An evaluation of knowledge management maturity', Expert Systems with Applications, 36(2), pp. 4087-4100.

Hsieh, P.J., Lin, B., Lin, C. and Chen, C. (2020) 'The evolution of Knowledge Navigator Model: The construction and application of KNM 2.0', Expert Systems with Applications, 148, 113209.

Hubert, C. and Lemons, D. (2019) APQC's Levels of Knowledge Management Maturity. Houston: APQC.

ISO (2018) ISO 30401:2018 Knowledge management systems - Requirements. Geneva: International Organization for Standardization.

McKinsey & Company (2025) The State of AI in 2025: Agents, Innovation, and Transformation. New York: McKinsey & Company.

Nonaka, I. and Takeuchi, H. (1995) The Knowledge-Creating Company. New York: Oxford University Press.

North, K. and Kumta, G. (2018) Knowledge Management: Value Creation Through Organizational Learning. 2nd edn. Cham: Springer.

Paulk, M.C., Curtis, B., Chrissis, M.B. and Weber, C.V. (1993) 'Capability maturity model for software, version 1.1', IEEE Software, 10(4), pp. 18-27.

Pee, L.G. and Kankanhalli, A. (2009) 'A model of organisational knowledge management maturity based on people, process, and technology', Journal of Information and Knowledge Management, 8(2), pp. 79-99.

Rios, J.C., Rodríguez-Elias, O.M. and Pino, F.J. (2011) 'A systematic review of software process maturity models', Software Quality Journal, 19(4), pp. 1-25.

Šajeva, S. and Jucevičius, R. (2010) 'The model of knowledge management system maturity and its application in Lithuanian organizations', Social Sciences, 3(69), pp. 57-68.

Thornley, C., Carcary, M., Connolly, N., O'Duffy, M. and Pierce, J. (2016) 'Developing a maturity model for knowledge management in the digital age', in Galbraith, B. and Moffett, S. (eds.) Proceedings of the 17th European Conference on Knowledge Management. Reading: Academic Conferences and Publishing International, pp. 874-880.

Yin, R.K. (2018) Case Study Research and Applications: Design and Methods. 6th edn. Thousand Oaks, CA: Sage.

The Role of Servant Leadership and Technology in Driving Organizational Citizenship Behaviour

Silvia Martelo-Landroguez, Jerónimo García-Fernández, Gabriel Cepeda-Carrión
Universidad de Sevilla, Seville, Spain

Abstract

The continuous change process in which today's society is immersed increasingly requires leaders to behave ethically and consider all stakeholders in the company (Al-Asadi et al., 2019). As a result, leadership has evolved from approaches that are more focused on how the personality of the leader influences individuals (e.g., transactional or transformational leadership) to more holistic approaches to leadership, such as servant leadership (SL), which is more focused on the individuals themselves (Panaccio et al., 2015).

Previous research has generally indicated a positive effect of SL on employee outcomes (Bavik, Bavik, and Tang, 2017; Lu, Falahat, and Cheah, 2024). However, more work is necessary to ascertain how and through what mechanisms SL influences such outcomes (Lu et al., 2024). In this regard, we have identified a fundamental element: technological adaptation. For many years, there has been a trend towards the digital transformation of companies, a process that has been accelerated by the COVID-19 pandemic situation. As the need for technological adaptation of companies is affecting most industries, the role that technology assumes for all parties in firms must be considered today.

This study examines the influence of SL and technological adaptation capability (TAC) on employees' organizational citizenship behaviour (OCB), using a sample of 53 leaders and 481 knowledge workers responsible for sports training activities that require advanced knowledge and empathy in fitness centres in Spain.

Our results indicate that establishing a TAC, by which knowledge workers encourage subordinates to make their own technological tool choices, is feasible and positively mediates the relationship between SL and OCB. The identification of SL as a catalyst for knowledge-sharing behaviour is largely consistent with the conclusions of authors such as Srivastava, Bartol and Locke (2006), who assert that servant leaders provide guidance and recognition to employees, thereby creating the context in organizations for sharing knowledge.

Keywords: Leadership, technology, employee outcomes, citizenship behaviour, knowledge sharing.

Paper type: Academic Research Paper

1 Introduction

Knowledge workers are highly qualified employees with highly specialized knowledge who become the primary driver of innovation within companies. As the environment change, they and their organizations must adapt to meet new challenges by keeping their team members motivated, confident, and aligned with the organization's goals through leadership (Alvesson, 2019).

The ongoing societal transformation process requires leaders to behave ethically and consider all stakeholders in the company (Al-Asadi et al., 2019). Traditional leadership styles fail to generate the intrinsic motivation that is currently leading to a higher level of organizational performance (Su et al., 2020). As a result, leadership has evolved from leader-focused approaches (e.g., transactional, or transformational leadership) to more holistic approaches, such as SL, which emphasizes individuals (Panaccio et al., 2015). Leaders must secure employee commitment by being the first to be committed to the organization (Djibo, Desiderio, and Price, 2010), focusing on follower development and achieving a sustainable competitive advantage that will allow organization survival in today's markets.

Previous research has generally indicated a positive effect of SL on employee outcomes (e.g., Bavik et al., 2017; Lu et al., 2024). Although OCB has been identified as a contributor to a more positive work environment, leading to increased organizational performance, lower turnover, and greater productivity, its underlying mechanisms are poorly understood (Koning and Van Kleef, 2015; Newman et al., 2017). The link between SL and OCB is still considered a black box (Addisu, Abebe, and Assefa, 2024). Several mediators have been proposed between SL and OCB, such as employee empathy, perceptions of leaders' ethics, perceived procedural justice and trust, and group service climate (Qiu and Dooley, 2022; Reiley and Jacobs, 2016). However, more work is necessary to ascertain how and through what mechanisms SL influences both organizational and individual-level outcomes (Lu et al., 2024). Servant leaders should cultivate an organizational culture that motivates employees to engage in OCB because they want to and not because they expect external rewards (Thomas, Ambrosini, and Hughes, 2019).

In this regard, TAC is key. For many years, there has been a trend towards the digital transformation of companies. COVID-19 pandemic situation has accelerated this process. Hence, the role that technology assumes for all parties in firms must be considered today. Suppose that employees perceive that leaders adopt technological systems to aid their job performance. In that case, they will experience higher levels of participation, which will motivate them to perform OCB. Specifically, our research question is: to what extent do SL and TAC significantly influence the OCB of employees? This aspect has been treated poorly in the management domain and the context of sports (Welty Peachey et al., 2015). Therefore, technology is incorporated into the proposed model to analyse whether SL can encourage its use and influence employee behaviour.

2 Theoretical Background

2.1 Organizational citizenship behaviour (OCB)

OCB has been conceptualized in many ways (Klotz et al., 2018). Williams and Anderson (1991) developed the most popular conceptualization, distinguishing between organization-focused OCB (i.e., actions aimed at the compliance with organizational rules and norms) and individually-focused OCB (i.e., actions explicitly aimed at helping individuals). This study uses such division into categories in terms of the target or direction of behaviour.

OCB includes behaviours that are of benefit to an organization but are outside the formal requirements of the job and rewards (Koning and Van Kleef, 2015). In this vein, Organ (1997) contends that employee behaviour does not need to be discretionary or voluntary to be considered an act of citizenship, it only needs to be contextual because it contributes to the organizational context supporting task performance. These behaviours are also referred to as extra-role behaviours, which are behaviours that cannot be required of employees for a particular job (Morrison, 1994). In any organization, at any given moment, any knowledge worker could do something that, albeit unrequired, could be helpful for other knowledge workers or the organization itself (Yam et al., 2017). Finding ways to encourage such behaviour in knowledge workers without directly requesting is essential.

2.2 Servant leadership (SL)

The growing level of knowledge in society (i.e., regarding people and employees) means that leadership styles associated with direct supervision are, in many cases, no longer effective and have bowed to other more successful approaches related to empowerment (Zu, 2019).

Greenleaf (1970) coined the term 50 years ago, and since then, numerous literature reviews have been conducted on SL (e.g., Parris and Welty Peachey, 2013; van Dierendonck, 2011). However, there is still a need to reach a consensus on its definition or theoretical framework (Mcquade, Harrison, and Tarbert, 2021). The seminal definition of SL proposed that: “The servant leader is servant first... it begins with the natural feeling that one wants to serve, to serve first. Then conscious choice brings one to aspire to lead” (Greenleaf, 1970, p. 7). We take the definition by Eva et al. (2019) as a reference, providing a solid basis for operationalizing the measures of SL that our analysis requires (Qiu and Dooley, 2022). We refer to SL as “an other-oriented approach to leadership, manifested through one-on-one prioritization of follower individual needs and interests, and outward reorienting of their concern for self towards concern for others within the organization and the larger community” (Eva et al., 2019, p. 114).

2.3 Technological adaptation capability (TAC)

The COVID-19 pandemic has changed our understanding of the world, affected the daily life of knowledge workers, and forced organizational structures to adapt to the latest technological formulas. Technology and its outputs (such as databases, files, and protocols) enable organizations to retain and make use of their valuable technology skills (Gold, Malhotra, and Segars, 2001). This study refers to technology as all the hardware and software that a company needs to use to achieve its goals (Laudon and Laudon, 2013). Technology can also be seen as all the products, methods, inventions, and standards that can be used to generate and share information (Kroenke, 2013) or as a tool for the management, storage, and dissemination of structured knowledge. Therefore, TAC refers to the firm’s ability to use technology to support work practices and improve working conditions (Ipe, 2003).

TAC constitutes an essential element of the structural dimension necessary for the mobilization of social capital to create new knowledge. Such capability is helpful because it supports better communication and collaboration, facilitates the organization and search for knowledge, and stimulates innovation (Alavi and

Leidner, 2001). If an organization aims to make its workers more willing to share what they know, it should first increase their interest in the use of technology (Tseng, 2017). Therefore, it should be kept in mind that if a firm wants to make its knowledge workers more willing to use technology, then the leader's involvement is critical.

3 Development of Hypotheses

As previously stated, SL refers to managing organizational challenges by subordinating the leader's personal interests to those of stakeholders (van Dierendonck, 2011). According to Otero-Neira, Varela-Neira and Bande (2016), SL promotes the development of employees as servants and the display of positive organizational behaviour.

It should be noted that if someone wants team members to develop certain behaviours, it must first be shown. One cannot ask someone to do something one would not do oneself (Cai et al., 2022). If knowledge workers showed OCB, their followers could understand its intrinsic value (Brown, Treviño, and Harrison, 2005), emulate their behaviours and engage in other servant behaviours. For example, they could share knowledge for the development of others (Tripathi et al., 2021).

There is a stream of studies related to the effects of SL on the relationship between employees and organizations (e.g., Addisu et al., 2024; Panaccio et al., 2015). Social learning theory and social exchange theory are mainly used to explain the relationship between these variables (Yang and Wei, 2018). Social learning theory (Bandura, 1977) states that individuals learn to behave in an appropriate manner through the observation of the behaviour of credible role models and imitating this behaviour. Under this framework, servant leaders are plausible models (Bavik et al., 2017; Tripathi et al., 2021). According to social exchange theory (Blau, 1964), once individuals benefit from the activities and actions of a firm, they are committed to seeking compensation and reciprocity. Social exchange theory suggests that individuals need to reciprocate when a servant leader cares about their needs. Therefore, we propose the following hypothesis:

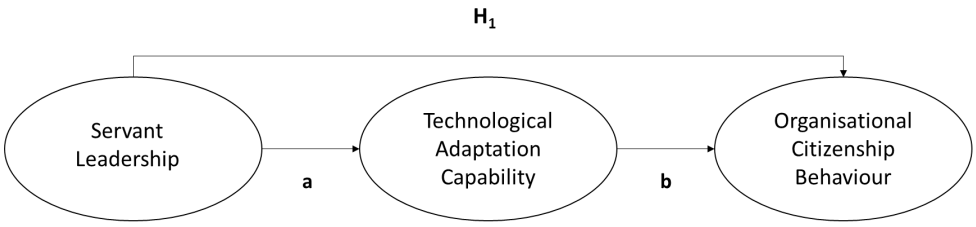
H1: SL positively influences OCB.

One of the consequences of the COVID-19 pandemic has been that people who had never used technologies have started to use them, motivated by necessity and social pressure (Batra, 2020). The importance of technology nowadays in organizations of any type requires the search for variables that promote technological transformation. In this regard, SL have been identified as a leadership style that positively influences technological adaptation (Hanif, Baloch, and Baig, 2020; Tseng, 2017). Therefore, TAC could strengthen the influence of SL on OCB by motivating knowledge workers to perform in the workplace (Zhou, Wang, and Breedlove, 2021). To the best of our knowledge, previous research focused only on how leaders can influence the adoption and implementation of technologies (Neufeld, Dong, and Higgins, 2007). The effect of TAC on employee behaviour has yet to be tested.

These arguments suggest that many organizations have used information technology to better cope with the challenges of operating in a turbulent environment (Sigala, 2003). The effective use and management of technology (i.e., TAC) constitute essential qualities for a company that could directly influence OCB and mediate the indirect relationship between SL and OCB. A possible explanation for this rests on the need to maintain knowledge within the organization and make it accessible to workers (Cepeda-Carrion et al., 2012). From this point of view, the leader leads by example using the company technologies, with technological adaptation of subordinates being a result of supervisor's SL and a mechanism by which servant leaders enhance employees' OCB. Therefore, we propose the second hypothesis:

H2: TAC positively mediates the relationship between SL and OCB.

The proposed relationships between the variables are shown in Figure 1.



H_2= Servant Leadership→ Technological Adaptation Capability → Organisational Citizenship Behaviour = a*b

Figure 1: Research model

4 Methodology

4.1 Research context

Fitness centres have been chosen as a suitable context due to its unique characteristics. Their management is a complicated task. Among the main errors, we would highlight: 1) needing more precise positioning and medium-term planning, 2) not using analysis and control tools, and 3) needing to work as a team. One way to overcome these problems is by a leadership style aligned with continuous changes in the organizational environment (Woo, 2018).

Fitness centres have adapted to COVID-19 pandemic by digitalizing their management and sports services (Batra, 2020). This digital transformation, while emphasized during the pandemic, constitutes a change over recent years (Jones, Ratten, and Hayduk, 2020). Fitness industry trends, demanding more excellent training and the applicability of technological skills in employees, influence this situation.

Fifty-three fitness centre leaders and 481 employees of 4 Spanish fitness chain participated in this study. Most of the respondents were women (56.7%), aged between 20 and 40 years (78.2%), and had a university degree (70.4%).

4.2 Measures

The executive SL scale proposed by Reed, Vidaver-Cohen and Colwell (2011), with 25 items, was chosen to measure SL. This scale includes the five Greenleaf's (1970) factors: interpersonal support, community building, altruism, egalitarianism, and moral integrity. The Lee and Allen's (2002) 16-item scale was chosen to measure the two target-based OCB dimensions: OCBO and OCBI. According to Yin and Yang (2010), TAC was measured by asking whether the information technology (IT) system is easy to use and helpful in performing the required tasks. Employees were also asked about the support of managers when using the IT system and about the availability of resources needed to use it.

4.3 Data analysis

We used PLS-SEM to test the proposed model since every measure was operationalized as composites (Cepeda-Carrion, Cegarra-Navarro, and Cillo, 2019). Our PLS-SEM analysis is explanatory and predictive (Henseler, 2018), focusing on testing the hypotheses and considering the predictability of indicators of the dependent variable (i.e., OCB). Model assessment with PLS-SEM involves the evaluation of the measurement and the structural model (Hair, Sarstedt, and Ringle, 2019). SmartPLS 3.3.9 was utilized to obtain results from the data analysis (Ringle, Wende, and Will, 2005). To prevent problems of common-method variance (CMV), the Kock and Lynn (2012) test of erratic collinearity was applied. This test is based on VIF values, whereby VIF values greater than 3.3 indicate collinearity problems. A maximum VIF of 1.04 was found, implying no CMV problems.

5 Results

5.1 Measurement model

We addressed individual reliability, which is confirmed by the fact that all constructs' standardized loadings exceed 0.7. Composite reliability measures are all higher than 0.8. The convergent validity threshold of 0.5 is surpassed by the average variance extracted (AVE) values. Both the criterion proposed by Fornell and Larcker (1981) and the HTMT ratio, provide evidence of discriminant validity (Henseler, Ringle, and Sarstedt, 2015).

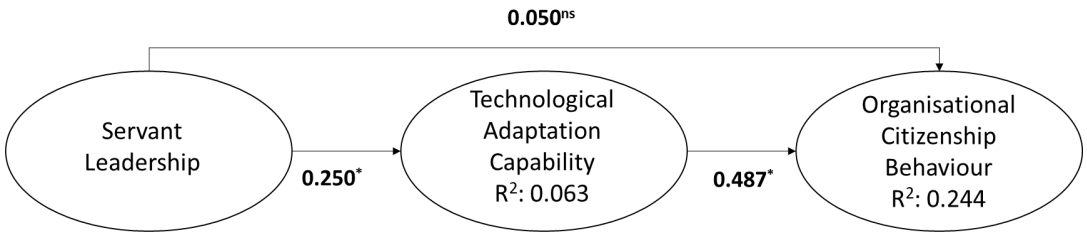
5.2 Structural model

After evaluating the absence of any collinearity issues between the constructs (Hair et al., 2019), the assessment of the path coefficients' sign, magnitude, and significance was carried out. Bootstrapping with 5,000 resamples was used to obtain confident intervals for evaluating the statistical significance of the path coefficients (Chin, 1998).

Our mediation hypothesis is analysed following Nitzl, Roldan and Cepeda's (2016) procedure. A full mediation of TAC was found in the relationship between SL and OCB. No control variables that influence our model were found. The percentile bootstrap at the 95% confidence interval was used to test hypotheses (Table 1). The effect size f^2 is reported, enabling the extent to which the explained variance of the dependent variables R^2 changes to be ascertained if a specified construct is absent from the model. Figure 2 shows the outcomes of the structural model.

Table 1: Structural Model (Bootstrapping)

Path	Path Coefficient	5%	95%	p-value	R^2	f^2
SL → OCB	0.050	0.029	0.149	0.182	0.244	0.003
SL → TAC	0.250	0.227	0.330	0.030	0.063	0.067
TAC → OCB	0.487	0.422	0.546	0.000	0.244	0.290
SL → TAC → OCB	0.122	0.108	0.163	0.030	0.244	0.357



* p<0.05; ns non-significant

Figure 2: Model results

R^2=0.244 confirmed model predictability for OCB. The PLSpredict procedure was also run, following Shmueli et al. (2019) (Table 2), which, in line with R^2, indicated high predictive power because all OCB dimensions exhibit lower residual errors for PLS-SEM than for the linear model (LM).

Table 2: PLSpredict

PLS			LM		PLS-LM	
	RMSE	Q^2_predict		RMSE		RMSE
OCBO	0.998	0.008	OCBO	0.999	OCBO	-0.001
OCBI	0.999	0.005	OCBI	1.003	OCBI	-0.006

6 Discussion

Knowledge workers can exercise a positive influence on how other people understand and interact with each other. To achieve this, it is essential to foster an attitude and mindset that is more receptive to new technology and the opinions of subordinates. In doing so, servant leaders play a crucial role in improving TAC while helping those affected by technological developments to adjust to these changes. One possible explanation of why SL is so essential for knowledge workers is that while only the leader's actions characterize other leadership styles, servant leaders are characterized by their willingness to serve others (Newman et al., 2017).

Organizations should encourage employees to display OCB due to its importance (Koning and Van Kleef, 2015). In this vein, our results support a direct and positive relationship between TAC and OCB. A possible explanation for this result would be that advances in the digital age are leading us towards new brain-technology interfaces that are changing our mental abilities and behaviours when interacting with others. For example, writing an email allows people to be more rational and control their emotions when expressing their ideas, especially when they are upset.

An interesting result is the nonsignificant direct relationship between SL and OCB. This circumstance invites us to believe that emotions, mixed feelings, and frustrations of not seeing themselves reciprocated can emerge among knowledge workers when their subordinates do not understand them or have different viewpoints. Our findings also confirm the critical role of TAC and support that it mediates the relationship between SL and OCB. It confirms what is said by authors such as Cepeda-Carrion et al. (2012), when they argue that since the various actors are forced to use a common language, TAC could deal with differences and agree on different views.

Based on the above, employees need internal motivation to show suitable behaviour. However, these behaviours cannot be easily acquired. This is the reason why organizations should find ways to encourage their employees to be good citizens (Khalili, 2017). Indeed, employee's engagement in OCB and how to encourage them to do so has been the subject of considerable research (Koning and Van Kleef, 2015). Empirical evidence suggests that leaders who serve their followers will create a more satisfied, committed, and empowered workforce with higher performance (Mayer, Bardes, and Piccolo, 2008). That is, employees will react to how their leader treats them. Therefore, servant leaders' care and commitment inspire employees to mimic this behaviour (Trong Tuan, 2017), facilitating the emergence of OCB.

6.1 Practical implications

First, we suggest that leaders must not only manage the environment in the organization and the employees directly but also be responsible for developing specific organizational capabilities. According to Tseng (2017), the leadership style is related to adaptability to technology and knowledge sharing. In this framework, knowledge workers will support the use of technology in the organization and enhance knowledge sharing among employees.

Second, TAC exerts an impact on the organizational environment, and it is, therefore, necessary to develop this capability among subordinates. Thus, the objective should be to design user-friendly information systems (i.e., allowing individuals with little experience to interact successfully with the system that influences employee satisfaction and motivation to perform better).

Third, the organizational environment will be improved by enhancing SL and TAC. Individuals are more likely to exhibit OCB in settings where suitable SL is combined with an appropriate TAC.

Fourth, managers can improve the TAC of their organizations through SL, which will result in OCB. For example, when leaders care about the employees and make their job easier, employees are more likely to be committed to the organization and perform for the benefit of the organization and other individuals because they want to, not just because they are asked to do so.

Furthermore, this change in employee behaviour and commitment will be a reward for the effort involved in the technological transformation. That could encourage the adaptation of this leadership style and the implementation of technology in more reluctant companies.

6.2 Limitations and future research

Future studies should be aligned to avoid the following limitations. First, there are self-report data for which concern arises regarding social desirability and the veracity of participant responses. Consistent with research suggesting that leaders overestimate their leadership abilities (Foster, Clarke, and Packard, 2018), future studies should use employee ratings of leadership. Second, the scales used in the study have been developed in the general business context and are relatively novel in the sports context. Future studies should combine organizations from different industries and countries to generalize the findings.

7 Conclusions

The first contribution of this research is to question existing models that manage the relationship between SL and OCB. Previous models have been applied from person to person. We now propose a person-to-person model through the means of technology. Our results indicate that establishing a TAC, whereby knowledge workers encourage subordinates to make their own technological tool choices, is possible and positively mediates the relationship between SL and OCB. The study also helps predict OCB through SL and TAC. This is largely consistent with the conclusions of authors such as Srivastava et al. (2006), who assert that servant leaders provide guidance and recognition to employees, thereby creating the context in organizations for sharing knowledge. Furthermore, our research model helps to predict OCB in innovative organizations such as fitness centres due to the increasing importance of technology in this sector, especially during COVID-19 lockdown.

8 Acknowledgements

This work was supported by Grant PID2021-127074NB-I00 funded by MICIU/AEI/10.13039/501100011033 and by the “European Union”.

9 Ethics declaration

Ethical clearance was not required for the research.

10 AI declaration

No AI tools were used for the creation of this paper.

References

Addisu, D.D., Abebe, E.A. and Assefa, T.T. (2024) Servant leadership and organizational citizenship behavior: The mediating role of perceived organizational politics and the moderating role of political skill in public service organizations, *Employee Responsibilities and Rights Journal*, pp 1-39.

Alavi, M. and Leidner, D.E. (2001) Review: Knowledge management and knowledge management systems: Conceptual foundations and research issues, *MIS Quarterly*, Vol. 25, No. 1, pp 107-136.

Al-Asadi, R., Muhammed, S., Abidi, O. and Dzenopoljac, V. (2019) Impact of servant leadership on intrinsic and extrinsic job satisfaction, *Leadership & Organization Development Journal*, Vol. 40, No. 4, pp 472-484.

Alvesson, M. (2019) Leadership: Convergence and divergence in leadership relations, *Journal of Management Inquiry*, Vol. 28, No. 3, pp 319-334.

Bandura, A. (1977) *Social Learning Theory*, Prentice-Hall, Englewood Cliffs.

Batra, D. (2020) The impact of the COVID-19 on organizational and information systems agility, *Information Systems Management*, Vol. 37, No. 4, pp 361-365.

Bavik, A., Bavik, Y.L. and Tang, P.M. (2017) Servant leadership, employee job crafting, and citizenship behaviors: A cross-level investigation, *Cornell Hospitality Quarterly*, Vol. 58, No. 4, pp 364-373.

Blau, P.M. (1964) *Exchange and Power in Social Life*, Wiley, New York.

Brown, M.E., Treviño, L.K. and Harrison, D.A. (2005) Ethical leadership: A social learning perspective for construct development and testing, *Organizational Behavior and Human Decision Processes*, Vol. 97, No. 2, pp 117-134.

Cai, D., Wang, H., Yao, L., Li, M. and Men, C. (2022) A relational identification perspective on why and when servant leadership foster employees' extra-role customer service, *Leadership & Organization Development Journal*, Vol. 43, No. 5, pp 705-718.

Cepeda-Carrion, G., Cegarra-Navarro, J.-G. and Cillo, V. (2019) Tips to use partial least squares structural equation modelling (PLS-SEM) in knowledge management, *Journal of Knowledge Management*, Vol. 23, No. 1, pp 67-89.

Cepeda-Carrion, G., Cegarra-Navarro, J.-G. and Jimenez-Jimenez, D. (2012) The effect of absorptive capacity on innovativeness: Context and information systems capability as catalysts, *British Journal of Management*, Vol. 23, No. 1, pp 110-129.

Chin, W. (1998) The partial least squares approach to structural equation modelling, *Modern Methods for Business Research*, Vol. 295, No. 2, pp 295-336.

Djibo, I.J., Desiderio, K.P. and Price, N.M. (2010) Examining the role of perceived leader behavior on temporary employees' organizational commitment and citizenship behavior, *Human Resource Development Quarterly*, Vol. 21, No. 4, pp 321-342.

Eva, N., Robin, M., Sendjaya, S., van Dierendonck, D. and Liden, R.C. (2019) Servant leadership: A systematic review and call for future research, *The Leadership Quarterly*, Vol. 30, No. 1, pp 111-132.

Fornell, C. and Larcker, D.F. (1981) Evaluating structural equation models with unobservable variables and measurement error, *Journal of Marketing Research*, Vol. 18, No. 1, pp 39-50.

Foster, C.A., Clarke, J.A. and Packard Jr, G.A. (2018) Can I see the real me? Leadership ability and the better-than-average effect, *Military Psychology*, Vol. 30, No. 5, pp 390-397.

Gold, A.H., Malhotra, A. and Segars, A.H. (2001) Knowledge management: An organizational capabilities perspective, *Journal of Management Information Systems*, Vol. 18, No. 1, pp 185-214.

Greenleaf, R.K. (1970). *The Servant as Leader*, The Robert K. Greenleaf Center, Indianapolis.

Hair, J.F., Sarstedt, M. and Ringle, C.M. (2019) Rethinking some of the rethinking of partial least squares, *European Journal of Marketing*, Vol. 53, No. 4, pp 566-584.

Hanif, M.I., Baloch, Z. and Baig, S. (2020) How servant leadership affect the organizational trust with mediating role of technological innovation? *International Review of Management and Marketing*, Vol. 10, No. 5, pp 74-84.

Henseler, J. (2018) Partial least squares path modeling: Quo vadis? *Quality and Quantity*, Vol. 52, No. 1, pp 1-8.

Henseler, J., Ringle, C.M. and Sarstedt, M. (2015) A new criterion for assessing discriminant validity in variance-based structural equation modelling, *Journal of the Academy of Marketing Science*, Vol. 43, No. 1, pp 115-135.

Ipe, M. (2003) Knowledge sharing on organizations: A conceptual framework, *Human Resource Development Review*, Vol. 2, No. 4, pp 337-359.

Jones, P., Ratten, V. and Hayduk III, T. (2020) Sport, fitness, and lifestyle entrepreneurship, *International Entrepreneurship and Management Journal*, Vol. 16, No. 3, pp 783-793.

Khalili, A. (2017) Transformational leadership and organizational citizenship behavior: The moderating role of emotional intelligence, *Leadership & Organization Development Journal*, Vol. 38, No. 7, pp 1004-1015.

Klotz, A.C., Bolino, M.C., Song, H. and Stornelli, J. (2018) Examining the nature, causes, and consequences of profiles of organizational citizenship behavior, *Journal of Organizational Behavior*, Vol. 39, No. 5, pp. 629-647.

Kock, N. and Lynn, G.S. (2012) Lateral collinearity and misleading results in variance-based SEM: An illustration and recommendations, *Journal of the Association for Information Systems*, Vol. 13, No. 7, pp 546-580.

Koning, L.F. and Van Kleef, G.A. (2015) How leaders' emotional displays shape followers' organizational citizenship behavior, *The Leadership Quarterly*, Vol. 26, No. 4, pp 489-501.

Kroenke, D. (2013) *Using MIS* (6th ed.), Prentice Hall, Upper Saddle River.

Laudon, K. and Laudon, J.P. (2013) *Management Information Systems: Managing the Digital Firm* (13th ed.). Prentice Hall, New York.

Lee, K. and Allen, N.J. (2002) Organizational citizenship behavior and workplace deviance: The role of affect and cognitions, *Journal of Applied Psychology*, Vol. 87, No. 1, pp 131-142.

Lu, J., Falahat, M. and Cheah, P.K. (2024) A systematic literature review on the relationship between servant leadership and its team and organizational level outcomes, *Journal of Organizational Change Management*, Vol. 37, No. 2, pp 255-282.

Mayer, D.M., Bardes, M. and Piccolo, R.F. (2008) Do servant-leaders help satisfy follower needs? An organizational justice perspective, *European Journal of Work and Organizational Psychology*, Vol. 17, No. 2, pp 180-197.

Mcquade, K.E., Harrison, C. and Tarbert, H. (2021) Systematically reviewing servant leadership, *European Business Review*, Vol. 33, No. 3, pp 465-490.

Morrison, E.W. (1994) Role definitions and organizational citizenship behavior: The importance of the employee's perspective, *Academy of Management Journal*, Vol. 37, pp 1543-1567.

Neufeld, D.J., Dong, L. and Higgins, C. (2007) Charismatic leadership and user acceptance of information technology, *European Journal of Information Systems*, Vol. 16, pp 494-510.

Newman, A., Schwarz, G., Cooper, B. and Sendjaya, S. (2017) How servant leadership influences organizational citizenship behavior: The roles of LMX, empowerment and proactive personality, *Journal of Business Ethics*, Vol. 145, No. 1, pp 49-62.

Nitzl, C., Roldan, J.L. and Cepeda, G. (2016) Mediation analysis in partial least squares path modelling: Helping researchers discuss more sophisticated models, *Industrial Management and Data Systems*, Vol. 116, No. 9, pp 1849-1864.

Organ, D.W. (1997) Organizational citizenship behavior: It's construct clean-up time, *Human Performance*, Vol. 10, pp 85-97.

Otero-Neira, C., Varela-Neira, C. and Bande, B. (2016) Supervisory servant leadership and employee's work role performance: A multilevel mediation model, *Leadership and Organization Development Journal*, Vol. 37, No. 7, pp 860-881.

Panaccio, A., Henderson, D.J., Liden, R.C., Wayne, S.J. and Cao, X. (2015) Toward an understanding of when and why servant leadership accounts for employee extra-role behaviors, *Journal of Business and Psychology*, Vol. 30, No. 4, pp. 657-675.

Parris, D.L. and Welty Peachey, J.W. (2013) A systematic literature review of servant leadership theory in organizational contexts, *Journal of Business Ethics*, Vol. 113, No. 3, pp 377-393.

Qiu, S. and Dooley, L. (2022) How servant leadership affects organizational citizenship behavior: The mediating roles of perceived procedural justice and trust, *Leadership & Organization Development Journal*, Vol. 43, No. 3, pp 350-369.

Reed, L., Vidaver-Cohen, D. and Colwell, S. (2011) A new scale to measure executive servant leadership: Development, analysis, and implications for research, *Journal of Business Ethics*, Vol. 101, No. 3, pp 415-434.

Reiley, P.J. and Jacobs, R.R. (2016) Ethics matter: moderating leaders' power use and followers' citizenship behaviors, *Journal of Business Ethics*, Vol. 134, pp 69-81.

Ringle, C.M., Wende, S. and Will, A. (2005) *SmartPLS 2.0*, University of Hamburg, Hamburg.

Shmueli, G., Sarstedt, M., Hair, J.F., Cheah, J.H., Ting, H., Vaithilingam, S. and Ringle, C.M. (2019) Predictive model assessment in PLS-SEM: Guidelines for using PLSpredict, *European Journal of Marketing*, Vol. 53, No. 11, pp 2322-2347.

Sigala, M. (2003) The information and communication technologies productivity impact on the UK hotel sector, *International Journal of Operations and Production Management*, Vol. 23, No. 10, pp 1224-1245.

Srivastava, A., Bartol, K.M. and Locke, E.A. (2006) Empowering leadership in management teams: Effects on knowledge sharing, efficacy, and performance, *Academy of Management Journal*, Vol. 49, No. 6, pp 1239-1251.

Su, W., Lyu, B., Chen, H. and Zhang, Y. (2020) How does servant leadership influence employees' service innovative behavior? The roles of intrinsic motivation and identification with the leader, *Baltic Journal of Management*, Vol. 15, No. 4, pp 571-586.

Thomas, L., Ambrosini, V. and Hughes, P. (2019) The role of organizational citizenship behaviour and rewards in strategy effectiveness, *The International Journal of Human Resource Management*, Vol. 30, No. 18, pp 2628-2660.

Tripathi, D., Priyadarshi, P., Kumar, P. and Kumar, S. (2021) Does servant leadership affect work role performance via knowledge sharing and psychological empowerment? V*INE Journal of Information and Knowledge Management System*, Vol. 51, No. 5, pp 792-812.

Trong Tuan, L. (2017) Knowledge sharing in public organizations: The roles of servant leadership and organizational citizenship behavior, *International Journal of Public Administration*, Vol. 40, No. 4, pp 361-373.

Tseng, S.M. (2017) Investigating the moderating effects of organizational culture and leadership style on IT-adoption and knowledge-sharing intention, *Journal of Enterprise Information Management*, Vol. 30, No. 4, pp 583-604.

Van Dierendonck, D. (2011) Servant leadership: A review and synthesis, *Journal of Management*, Vol. 37, No. 4, pp 1228-1261.

Welty Peachey, J.W., Zhou, Y., Damon, Z.J. and Burton, L.J. (2015) Forty years of leadership research in sport management: A review, synthesis, and conceptual framework, *Journal of Sport Management*, Vol. 29, No. 5, pp 570-587.

Williams, L.J. and Anderson, S.E. (1991) Job satisfaction and organizational commitment as predictors of organizational citizenship and in-role behaviors, *Journal of Management*, Vol. 17, pp 601-617.

Woo, B. (2018) Relationship between servant leadership attributes and trust in leaders: A case of sport instructors in South Korea, *Sport Journal*, Vol. 20.

Wu, J., Liden, R.C., Liao, C. and Wayne, S.J. (2021) Does manager servant leadership lead to follower serving behaviors? It depends on follower self-interest, *Journal of Applied Psychology*, Vol. 106, No. 1, pp 152-167.

Yam, K.C., Klotz, A.C., He, W. and Reynolds, S.J. (2017) From good soldiers to psychologically entitled: Examining when and why citizenship behavior leads to deviance, *Academy of Management Journal*, Vol. 60, No. 1, pp 373-396.

Yang, Q.I. and Wei, H. (2018) The impact of ethical leadership on organizational citizenship behavior: The moderating role of workplace ostracism, *Leadership & Organization Development Journal*, Vol. 39, No. 1, pp 100-113.

Yin, G. and Yang, B. (2010) The construction of firm's IT capability and its impact on IT assimilation: An empirical investigation in China, *PACIS 2010 Proceedings*, p 74.

Zhou, L., Wang, L. and Breedlove, J. (2021) Impact of digital technology on optimizing organizational and social dynamics of the sport industry in China. In L. Zhou, L. Wang, & J. Breedlove (Eds.). *International Sport Business Management* (pp.145-168). Routledge.

Zu, L. (2019) Purpose-driven leadership for sustainable business: From the perspective of Taoism, *International Journal of Corporate Social Responsibility*, Vol. 4, No. 1, pp 1-31.

Sociodigital and Cultural Skills as Mediators of Technology Adoption in Higher Education Learning Ecosystems

Ana Lidia Quintero Ramírez[1], Gabriela Citlalli López-Torres[2], Silvia Alexandra Peña Gonzalez[3], Hannia Guadalupe García Vega[4], Francisco Javier Álvarez-Torres[4]

[1] National Polytechnic Institute, Mexico
[2] University of Aguascalientes, Mexico
[3] University of Cuenca, Ecuador
[4] University of Guanajuato, Mexico

Abstract

Technology adoption in higher education (HE) constitutes a complex process that goes beyond access to digital tools, this process involves individual competences, academic practices and sociodigital and cultural skills. This study focused on students in higher education from two countries in Latin America: Mexico and Ecuador, considering their paper of the digital and sociocultural skills in the development of the significant learning practices in the educative multicultural ecosystems.

For the analysis of the results, an exploratory design of mixed methods was used, with the application of a questionnaire with a Likert scale, who consist in a 150 sample, which one 100 was from Mexico and 50 from Ecuador. The data analysis was carried out in two stages, firstly, started with the descriptive analysis and the instrument reliability with Statgraphics® software, evaluating the internal consistency and the validity of analytical dimensions. Secondly, Structural Equation Modelling was employed based on partial minimum square (PLS-SEM) using SmartPLS® 4, with the principal objective of examining the structural relationships between the four dimensions and the capability of the model. The analytical model consists in four dimensions: digital and sociocultural skills, academic use of technology, ethical and responsible use of technology, and institutional conditions. For the analysis of the results, it used a conceptual model of technology which show four directional relationships between the constructs, such as, 1) Digital and sociocultural skills-academic use of technology, 2) Institutional conditions-academic use of technology, 3) Academic use of technology-ethical and 4) Responsible use of technology and digital and sociocultural skills- ethical and responsible use of technology. In this model, technology enhances learning and ethical practices from a transversal STEM perspective.

The results show that the digital and sociocultural skills impact significantly in the academic use of technology in both contexts, while the academic use of technology emerge as the principal predictor of their ethic and responsible use. In contrast, the direct relationship between the sociocultural skills and the ethical and responsible use of technology is not meaningful, which confirms the mediator role of the use of technology. Empirical evidence shows that the technology adoption in higher education must understanding as a mediated process with oriented learning and conditioning by institutional context.

Keywords: Technology adoption, sociodigital skills, Knowledge management, learning ecosystems

Paper type: Academic Research Paper

1 Introduction

The digital transformation in higher education (HE) is experiencing an unprecedented phenomenon marked by the convergence of Artificial Intelligence (AI), demographic changes, and the redefinition of job skills. This process goes beyond the incorporation of emerging technologies, as it raises the need for a thorough review of the ways in which learning, knowledge and skill formation are conceived, questioning not only how, but also why and what is learned in contemporary digital environments (European Commission, Joint Research Centre, 2022).

International frameworks such as DigComp 2.2 have reframed digital transformation as a people-centered process rather than merely technological infrastructure. From this perspective, digital competence involves the critical and responsible use of technology through the integration of knowledge, skills, and attitudes developed through educational processes (Vuorikari, Kluzer, Punie, 2022).The emphasis on skills training is justified by the need to thrive in an unstable environment, where, in the world of "infodemia" and misinformation, HE must train professionals capable of assessing the credibility and reliability of data sources. In this sense, the DigComp framework allows students to identify their own digital gaps and seek opportunities for self-development to stay updated with constant technological evolution (Cosgrove, Cachia, 2025).

Farias-Gaytan (2023) identifies a significant gap between the incorporation of technological tools and the effective development of digital skills in higher education institutions. In this context, digital transformation involves not only the digitalization of educational activities, but also cultural, organizational, and pedagogical changes aimed at strengthening advanced digital competences. Despite increasing interest in educational technologies, many academic programs still lack systematic strategies to develop digital literacy and promote the effective pedagogical integration of technology among teachers and students (Macías-Galeas, 2024).

In this context, digital transformation in HE is a process that transcends the incorporation of technologies, being deeply conditioned by the institutional culture, and the learning ecosystems in which it develops. In this framework, characterized by cultural diversity and institutional heterogeneity, technological adoption is configured as a complex phenomenon in which pedagogical practices, digital infrastructures and socio-digital skills focused on critical use interact, ethical and collaborative of knowledge interact (Macías-Galeas, 2024). However, a gap remains in empirical evidence about how these cultural and educational dimensions mediate technological adoption in multicultural learning ecosystems, particularly in HE. In response to this gap, the present research analyzes the adoption of digital technologies in university students from the perspective of socio-digital and cultural skills, so, to understand this process as a contextualized educational practice oriented, towards creation and management of knowledge.

2 Justification of the Research

Nowadays, there is a need to deepen approaches that support technological adoption in higher education, due to the accelerated growth of digital transformation. While digital transformation has promoted a broad incorporation of technologies in educational processes, there remains a limited understanding of the role that socio-digital and cultural skills play in the meaningful adoption of such technologies within complex learning ecosystems.

This study is justified by the growing need to understand technological adoption in higher education beyond the simple incorporation of digital tools. Although digital transformation has expanded the use of technologies in educational processes, there is still limited understanding of how socio-digital and cultural skills influence the meaningful, critical, and responsible use of technology within complex learning ecosystems. From a knowledge management perspective, these skills are essential for the creation, validation, and application of knowledge in digital environments.

From an empirical perspective, there is limited research analyzing the relationship between technological adoption, socio-digital competences, and cultural dimensions in higher education, particularly in Latin America. In this sense, the comparison between Mexico and Ecuador provides contextualized evidence within multicultural educational ecosystems. Furthermore, the study contributes to the development of educational strategies and institutional policies aimed at strengthening socio-digital skills and promoting meaningful and ethical technology use in higher education.

3 Research Problem Statement

In today's education environment, the need for digital transformation has become highly necessary, mainly because of continued technological progress and its impact on society. In 2024, the World Economic Forum revealed that 87% of companies anticipate using digital transformation in their processes, but only half believe they are prepared. This disparity underscores the urgent need for institutions, particularly higher-level institutions, to reconsider their operating models to remain relevant. According to UNESCO, digital innovation can support and improve education, helping to accelerate progress toward universal access to quality learning, as outlined in Sustainable Development Goal 4. Digital transformation in higher-level institutions involves the incorporation of new technologies and the transformation of practices, business models, and processes. This is directly related to the development of new and improved pedagogical approaches (Bucăţa et al., 2022).

Higher education institutions have increasingly adopted educational technologies such as digital platforms, e-learning systems, and AI-based tools. However, technology adoption does not necessarily guarantee the development of critical, transferable, and ethical digital competences among students. Existing research suggests that meaningful technology use depends on socio-digital skills, pedagogical integration, and institutional conditions that mediate the relationship between technology and learning processes (Feng et al., 2025). In addition, several institutions, particularly in developing countries, continue to face barriers related to digital infrastructure, connectivity, and digital literacy gaps among teachers and students (Alenezi and Akour, 2023).

In this context, the present study analyzes the adoption of digital technologies among higher education students in Mexico and Ecuador, considering these settings as multicultural learning ecosystems shaped by

sociocultural and institutional diversity. The study approaches technological adoption as a mediated process influenced by socio-digital and cultural skills that support the critical, ethical, and meaningful use of technology in learning and knowledge management processes.

4 Research Objectives

The general objective of this research is to analyze technological adoption in higher education students in two countries: Mexico and Ecuador, considering the role of socio-digital and cultural skills in developing meaningful learning practices within multicultural educational ecosystems.

5 Theoretical framework

Today's society lives in what is known as a "digital age", where the products and technological practices can quickly become essential parts of everyday life. These changes have impacted on how people communicate, consume information, and organize their life. These have repercussions in how learning is performed, how knowledge is created, and how people expect to be taught. This makes sense in the educational sector, who is reinvented under a technological light, where digital systems and networks are integral to the management and operation of modern institutions. This has obvious consequences for schools, universities, and other educational institutions (Selwyn, 2016).

5.1. Technological adoption in higher education

In the XX century, knowledge was seen as a final product or result of education, most of the time like a change in the conductivist conception of learning. This notion of learning as a product remains a common-sense way of understanding. Many students (and some educators) continue to view learning as largely acquiring knowledge and filling empty containers or ideas that Bereiter (2002) described as the "popular" theories of learning.

However, alternative perspectives have increasingly challenged this notion of learning. In a seminal study, Säljö (1979) identified different ways in which adults understand learning:

- Learning as a quantitative increase in knowledge
- Learning as memorization
- Learning as the acquisition of facts, skills and methods that can be retained and used as needed.
- Learning having meaning or abstract the meaning, that is, learning that involves relating parts of the subject to each other and with the real world.
- Learning how to interpret and understand reality in a different way

These two latter descriptions of learning as a continuous process highlight the role of prior experiences, reflection, and contextual interaction on knowledge construction. This conception of learning provides a perspective where digital technologies acquire educational value not through their mere presence, but through their integration into academic practices that support meaningful learning, problem solving, and knowledge application. Consequently, the use of technology for academic purposes becomes a central mechanism through which learning processes are mediated, and through which ethical and responsible practices can emerge.

From this perspective, technological adoption in higher education is understood as a mediated and context-dependent process, rather than a direct consequence of digital skills alone. Socio-digital and cultural competencies provide a necessary foundation for their integration into academic practice that meaningful and ethical technology use emerges (Kolmos, 2017; OECD, 2021). In this sense, the academic use of technology operates as a central mediating mechanism through which individual skills are transformed into ethical and responsible technology use. This view is closely aligned with transversal STEM educational approaches, which emphasize problem solving, interdisciplinary learning, and the practical application of knowledge as core learning processes. Within such approaches, technology functions not merely as a tool, but to support reflective and responsible engagement with real-world challenges. This process is further shaped by institutional conditions that enable or constrain opportunities for pedagogically meaningful engagement with technology, reinforcing the need to understand technological adoption within educational ecosystems rather than as an individual or purely technical phenomenon (Schiuma, 2011; Kozma, 2011)

Based on the theoretical perspective, technological adoption in higher education can be analytically structured through four interrelated dimensions that reflect the interplay between individual competencies, academic practices, ethical outcomes and institutional context. First, digital and sociocultural skills represent the foundational competencies that enable students to critically engage with digital technologies, integrating technical abilities with cognitive and social awareness (UNESCO, 2018; Redecker, 2017). Second, the academic

use of technology captures the extent to which digital tools are embedded into meaningful learning processes, supporting problem solving, knowledge construction, and application, particularly within STEM-oriented and practice-based learning environments (Kolmos, 2017; OECD, 2021). Third, ethical and responsible use of technology is conceptualized as a formative outcome that emerges through sustained academic engagement, reflecting students' capacity to use digital technologies in a reflective, responsible, and socially aware manner (Selwyn, 2016; Floridi et al., 2018). Finally, institutional conditions represent the contextual dimension that shapes access, opportunities, and pedagogical support for technology integration, highlighting the role of educational ecosystems in enabling or constraining technological adoption (Schiuma, 2011; Kozma, 2011). Together, these four dimensions provide a coherent framework for understanding technological adoption as a mediated process in which competencies, practices, ethics, and context interact dynamically.

5.2. Integrative conceptual model

Building on the four analytical dimensions identified, this study proposes an integrative conceptual model to explain the technological adoption in higher education. Considering the socio-constructivist and critical perspectives on educational technology, a model emerges that assumes that educational value of digital technologies is produced through their integration into academic practices and is shaped by broader institutional environments. This suggests that technological adoption in higher education is a mediated and context-dependent process, rather than a direct result of digital skills.

Contemplating the ethical and responsible use of technology, this is conceived as an outcome that arises from sustained academic engagement rather than as an attribute inherent to people with digital skills. So, in this context, ethical practices are framed as a formative result of technology-mediated learning experiences supported by institutional contexts, rather than as a direct consequence of technical competence alone. Together, this conceptual model integrates socio-digital skills, the academic and ethical use of technology, as well as institutional conditions, providing a coherent theoretical basis for analyzing technological adoption in higher education as a dynamic process shaped by individual competencies, academic practices and institutional ecosystems.

6 Methodology

Given the exploratory nature of the research and the complexity of the phenomenon under investigation, this study adopts an exploratory mixed-method research design, to capture contextual dynamics and actors' perspectives. Data was collected through an online questionnaire administered using Google Forms, which combined open-ended questions with a Likert scale item. The instrument was designed to explore students' experiences, perceptions, and practices related to technology use in academic contexts. The questionnaire was structured around four analytical dimensions derived from the literature and the theoretical framework:

- Digital and socio cultural skills
- Academic use of technology
- Ethical and responsible use of technology
- Institutional conditions

These dimensions enable a structured interpretation of the empirical material and support systematic analysis of students' technological adoption processes. The sample consisted of 150 students enrolled in university institutions from two Latin American countries: Mexico and Ecuador. Specifically, data were collected from 100 students in Mexico, from the University of Aguascalientes, University of Guanajuato and National Polytechnic Institute, and 50 students in Ecuador from the University of Cuenca. The sampling strategy was non-probabilistic and based on convenience, as participants were selected according to accessibility and institutional collaboration. This approach is consistent with exploratory research designs aimed at identifying structural patterns rather than achieving statistical representativeness.

The analysis was guided by an integrative analytical framework derived from the literature on digital competences, pedagogical ICT integration, digital sustainability, and 21st century skills. The four analytical dimensions provided a structured basis for interpreting the empirical material and analysing students' technological adoption processes. In particular, the dimension of academic use of technology was informed by UNESCO's digital competence frameworks and STEM-oriented educational approaches, conceptualizing technology use to support problem solving, interdisciplinary learning, and the practical application of knowledge.

Data analysis was conducted using Statgraphics® and SmartPLS 4 software, the first software analyzed the Likert scale items to summarize students´ responses and identify general patterns related to the four analytical

dimensions. SmartPLS 4 was employed for the exploration analysis of relationships among the dimensions, supported by the examination of technological adoption as a multicultural construct.

7 Analysis and Results

The results obtained from the four dimensions' analysis were interpreted as a structured process of technological adoption rather than isolated statical associations. This analysis integrates descriptive findings obtained through statistical analysis, emphasizing contextual similarities and differences between Mexico and Ecuador. Figure 1 represents the structural model of technological adoption of this research, which is integrated into four dimensions represented in blue circles and their respective indicators, represented by yellow rectangles. The directional relationships between these constructs illustrate a sequential and mediated process. Digital and sociocultural skills and institutional conditions act as antecedents that enable academic use of technology, which in turn functions as the central mechanism leading to ethical and responsible use. While a direct relationship between skills and ethical use is included, the model highlights that ethical practices primarily emerge through academic engagement. Overall, the model reflects a process-oriented and context-dependent interpretation of technological adoption, consistent with transversal STEM educational approaches.

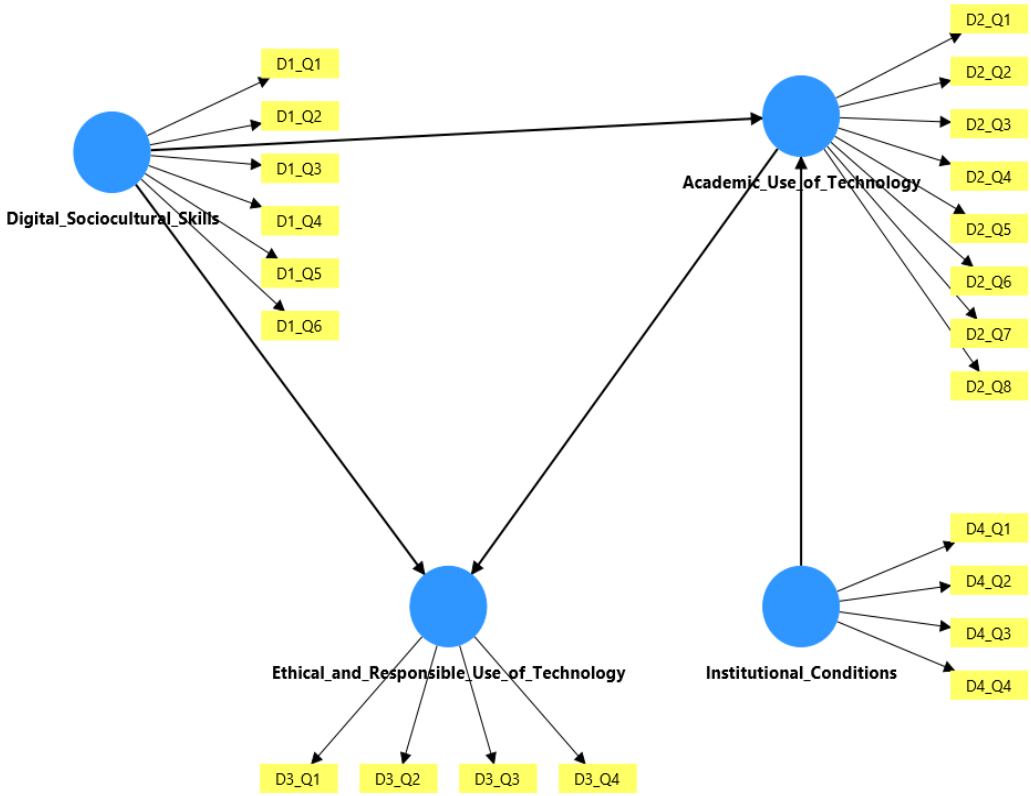


Figure 1: Conceptual model of technology-enhanced learning and ethical practices from a transversal STEM perspective. Source: SmartPLS 4.

The conceptual model proposes four directional relationships that collectively describe technological adoption in higher education as a mediated and context-dependent process. Rather than assuming linear or purely skill-based effects, the model emphasizes the interaction between individual competencies, academic practices, and institutional environments to shape meaningful and ethical use of technology. This model shows the four directional relationships between the constructs:

- Digital and sociocultural skills --- > Academic use of technology
  - This construct indicates that students' digital and sociocultural skills directly influence their ability to integrate technology into meaningful academic practices.
- Institutional conditions --- > Academic use of technology
  - This construct reflects the facilitating role of the institutional environment in the academic use of technology, highlighting the importance of the educational context as an enabler of digital learning.

- Academic use of technology --- > Ethical and responsible use of technology
    - This relationship makes the academic use of technology a central mechanism, suggesting that digital ethics emerge from structured learning experiences.
- Digital and sociocultural skills --- > Ethical and responsible use of technology
    - This direct relationship represents an alternative path between skills and ethics of the responsible use of technology.

Prior to examining the structural relationships among constructs, the reliability of the measurement instrument was assessed to ensure the robustness of the analytical dimensions. This analysis indicates acceptable internal consistency across all dimensions in both samples (Mexico and Ecuador). The table 1 shows the Cronbach's alpha values who meet or exceed recommended thresholds for exploratory research. Convergent validity, assessed through Average Variance Extracted (AVE), is satisfactory for most constructs. Although the AVE value for digital and sociocultural skills in the Ecuadorian sample is slightly below the conventional threshold, it remains acceptable within an exploratory and cross-national research design.

Table 1**:** Reliability and Average Variance Extracted (AVE) of the dimensions

Dimensions	**Items**	**Mexico**		**Ecuador**	
		Cronbach's α	**AVE**	**Cronbach's α**	**AVE**
Digital and sociocultural skills	6	0,807	0,506	0,697	0,430
Academic use of technology	8	0,906	0,605	0,874	0,540
Ethical and responsible use of technology	4	0,831	0,660	0,771	0,587
Institutional conditions	4	0,702	0,512	0,748	0,553

Source: Own elaboration

Once the reliability of the measurement instrument was determined, determination coefficient is considered, which considers the determination of a dependent variable, which is explained by the set predictors of the model (Kuehl, 2000). The results in Mexico indicate that the model explains a substantial proportion of variability in the endogens constructs, the use of technology (R^2= 0,572) and ethical and responsible use of technology (R^2= 0,623). For the case of Ecuador, the academic use of technology shows a R^2= 0,648 and for the ethical and responsible use of technology 0,629, this is significant superior for the dimension 2 in the case of Mexico.

Although the magnitude of some relationships varies between Mexico and Ecuador, the overall structure of the model remains consistent with the technological adoption, which follows a similar formative process within multicultural higher education ecosystems, despite contextual differences. The table 2 shows the structural relationship between the dimensions of the technological adoption in Mexico and Ecuador.

Table 2**:** Structural relationships of the technological adoption model across Mexico and Ecuador.

Construct	**Mexico**			**Ecuador**		
	Route coefficient (β)	**p-value**	**t-value**	**Route coefficient (β)**	**p-value**	**t-value**
Digital and sociocultural skills ---> Academic use of technology	0,569	0,000	8,712	0,686	0,000	6,969
Institutional conditions ---> Academic use of technology	0,310	0,000	4,335	0,202	0,072	1,800
Academic use of technology ---> Ethical and responsible use of technology	0,761	0,000	10,041	0,638	0,000	5,171
Digital and sociocultural skills ---> Ethical and responsible use of technology	0,040	0,680	0,413	0,187	0,164	1,392

Source: Own elaboration

The results of the relationship between the dimensions show a consistent patron of technological adoption in both national contexts, although with variations in some of the constructs. For first instance, the digital and sociocultural skills revel a positive effect on the use of academic technology in Mexico (β= 0,569, $p < 0.05$) and Ecuador (β= 0,686, $p < 0.05$), who confirms the central role of the individual competences in the integration of technology in learning practices. This result confirms the literature who highlights the digital alphabetization and

sociocultural dimension as necessary conditions for a pedagogical use in technologies in universities (UNESCO, 2018; Redecker, 2017).

The construct of academic use of technology, emerges as the principal predictor of the ethical and responsible use of technology in both countries with the highest coefficients in México ($\beta = 0.761$, $p < 0.05$) and Ecuador ($\beta = 0.638$, $p < 0.05$). This supports the idea that the ethical practices in the use of technology doesn't develop in isolation, but build through structured academic experiences focused on learning, problem solving and knowledge application, in line with STEM and meaningful learning approaches ((Kolmos, 2017; OECD, 2021).

Empirical evidence revels that the institutional conditions represent an influence in the academic use of technology in Mexico, while in Ecuador ($\beta = 0{,}202$, $p = 0{,}072$), the relationship is weak and is not enough to achieve the level of conventional statistical significance. This difference suggests that, although the institutional conditions act as a facilitator of academic use, their influence depends on the specific educational context. Enhancing the fact that the technology adoption must understand as a perspective of the educational ecosystems more than a homogenous process (Schiuma, 2011; Kozma, 2011).

In contrast, the digital and sociocultural skills and the ethical and responsible use of technology didn't result as statistically significant in Mexico and Ecuador. This reinforces the idea that the digital competences, by his own, doesn't guarantee responsible practices, although these competences must work through the academic use of technology for generate ethical behaviours. Floridi, et al. (2028) highlights the importance of have a mediator of the academic use of technology, pointing it out as the central mechanism who articulates the individual competences, institutional context and formative results.

The qualitative analysis of results shows that, both in Mexico and in Ecuador, technological adoption is perceived as a process that goes beyond the technical domain of digital tools. Students' narratives confirm that academic use of technology is the key element that transforms digital skills into meaningful learning experiences and ethical, responsible practices. This perception highlights that technological value is not derived from mere access or technical competence, but from the ways in which digital tools are incorporated into academic activities such as problem-solving, research, collaboration, and applied learning. In both contexts, students associate the ethical and responsible use of technology with situations where technology is deliberately integrated into learning tasks and guided by academic expectations.

At the same time, narratives reveal that institutional conditions - such as availability of infrastructure, access to specialized software, and pedagogical support - shape the depth and quality of this academic commitment. which contributes to the contextual differences between Mexico and Ecuador. In general, these perspectives reinforce the conceptualization of technological adoption as a mediated and ecosystem-dependent process, aligned with cross-cutting STEM approaches that emphasize learning by doing, responsible reflection and engagement with real-world challenges.

8 Conclusions

The main objective of this study was to analyze technological adoption in higher education (HE) students from Mexico and Ecuador, considering the role of digital and sociocultural skills, academic and ethical use of technology, as well as institutional conditions in the development of meaningful learning practices within multicultural educational ecosystems. Based on the integrative conceptual model and a PLS-SEM analysis, the results identified that technological adoption is not a linear process or a direct result of the possession of digital skills, but a mediated, contextual and fundamentally formative phenomenon.

The results show that digital and sociocultural skills form the necessary basis for academic use of technology in both national contexts. However, these skills do not automatically translate into ethical and responsible practices. In this sense, the academic use of technology emerges as the central mechanism that links individual competencies with formative outcomes, confirming that digital ethics is built through structured learning experiences oriented to the pedagogical use of technology, and not as a direct consequence of technical mastery.

Also, the results show that institutional conditions play a relevant role in technological adoption, although their influence varies between Mexico and Ecuador. This variability reinforces the relevance of analyzing technological adoption from an educational ecosystem perspective, in which institutions act as enablers or limiters of academic use of technology, without altering the overall structure of the identified adoption process. Considering the STEM approach of transversal effect, the results underline the importance of conceiving technology as a means for active learning, problem-solving, and application of knowledge in significant academic contexts. In this framework, the ethical and responsible use of technology is not conceived as an isolated competition, but as an emerging result of educational practices that promote reflection, informed decision-making, and connection to real issues.

Finally, this study provides empirical and conceptual evidence to understand how technological adoption is generated in higher education as a process mediated by the academic use of technology and conditioned by the institutional context. These findings have relevant implications for higher education institutions, pointing to the need to strengthen pedagogical and organizational strategies that intentionally integrate technology into the curriculum, promoting meaningful digital learning and ethical practices within multicultural educational ecosystems.

9 Ethics declaration

This study was conducted in accordance with the ethical standards for participation of higher-level students. The data were collected through a voluntary questionnaire administered to higher education students in Mexico and Ecuador. All participants were informed about the purpose of the study and their right to withdraw at any time without consequences. Participation was anonymous, and no personally identifiable information was collected. The data was used exclusively for academic and research purposes, ensuring confidentiality and responsible handling of the data throughout the research process.

10 AI declaration

Artificial intelligence tools used specifically to assist with language refinement, clarity of expression, and structural editing of the text. The use of AI did not involve data analysis, generation of results, or interpretation of empirical findings. All theoretical content, methodological decisions, data analyses and conclusions remain the sole responsibility of the authors.

References

Alenezi, M. and Akour, M. (2023) Digital transformation blueprint in higher education: A case study of PSU, Sustainability, 15, 8204. https://doi.org/10.3390/su15108204

Bereiter, C. (2002) Education and mind in the knowledge age. Mahwah, NJ: Lawrence Erlbaum Associates.

Bucăţa, G., Popescu, F. and Tileagă, C. (2022) Digital transformation of higher education system, International Conference on Knowledge Based Organization, 28(1), pp. 158–168. https://doi.org/10.2478/kbo-2022-0025

Cosgrove, J. and Cachia, R. (2025) DigComp 3.0: European Digital Competence Framework – Fifth Edition, Publications Office of the European Union, Luxembourg. doi:10.2760/0001149.

European Commission, Joint Research Centre (2022) DigComp 2.2: The Digital Competence Framework for Citizens. With new examples of knowledge, skills and attitudes, Publications Office of the European Union, Luxembourg. doi:10.2760/115376.

Farias-Gaytan, et al., (2003) Digital transformation and digital literacy in the context of complexity within higher education institutions: a systematic literature review, https://doi.org/10.1057/s41599-023-01875-9

Feng, J., Yu, B., Tan, W. H., Dai, Z. and Li, Z. (2025) Key factors influencing educational technology adoption in higher education: A systematic review, PLOS Digital Health, 4(4), e0000764. https://doi.org/10.1371/journal.pdig.0000764

Floridi, L., Cowls, J., Beltrametti, M., et al. (2018). AI4People—An ethical framework for a good AI society. Minds and Machines, 28(4), 689–707

Kolmos, A. (2017). Problem-based and project-based learning in engineering education. Cambridge Handbook of Engineering Education Research.

Kozma, R. B. (2011). Transforming education: The power of ICT policies. UNESCO.

Kuehl, R. (2000). Design of experiments Statistical design principles and research analysis. Second edition, Thomson Learning, ISBN-0-534-36834-4.

Macías-Galeas, I. (2024) Transformation and Challenges Of Higher Education in the Digital Age: strategies for skill development in the 21st century YUYAY: Estrategias, Metodologías & Didácticas Educativas, Vol. 3, No. 1, pp. 17–33. doi:10.59343/yuyay.v3i1.57.

OECD. (2021). OECD Digital Education Outlook. OECD Publishing.

Redecker, C. (2017). European Framework for the Digital Competence of Educators (DigCompEdu). European Commission.

Säljö, R. (1979) Learning in the learner's perspective I: Some common-sense conceptions. Report No. 76. Gothenburg: Institute of Education, University of Gothenburg.

Schiuma, G. (2011). Managing knowledge assets and business value creation in organizations. IGI Global.

Selwyn, N. (2016) Education and technology: Key issues and debates. 2nd edn. London: Bloomsbury Academic

Singun, A. (2025) Unveiling the barriers to digital transformation in higher education institutions: a systematic literature review, Discover Education. https://doi.org/10.1007/s44217-025-00430-9

UNESCO. (2018). Digital skills critical for jobs and social inclusion. UNESCO.

Vuorikari, R., Kluzer, S. and Punie, Y. (2022) DigComp 2.2: The Digital Competence Framework for Citizens, EUR 31006 EN, Publications Office of the European Union, Luxembourg.

World Economic Forum (2024) Accelerating digital transformation for long-term growth, World Economic Forum. Available at: https://initiatives.weforum.org/digital-transformation/home

Beyond Profit, Through Data: How AI-Driven Academic Spin-offs Enact Public Value in Knowledge Ecosystems

Silvana Revellino
University of Salerno, Fisciano, Italy

Abstract

Artificial intelligence (AI) is increasingly positioned as a key driver of digital transformation, enabling organizations to enhance performance, decision-making, innovation and stakeholder engagement. However, for public, hybrid, and mission-driven organizations, the critical issue is not only how AI improves efficiency, but how it contributes to value creation beyond profit (Moore, 1995; 2013; Battilana and Lee, 2014). This paper investigates how AI-driven academic spin-offs enact public value within digital knowledge ecosystems.

Drawing on public value theory (Moore, 1995; 2013), the sociology of valuation (Stark, 2009; Dussauge et al., 2015; Kornberger et al., 2015), knowledge management literature (Davenport and Prusak, 1998; Zack, 1999), and research on AI and entrepreneurship (Shepherd and Majchrzak, 2022; Fossen et al., 2024), the study conceptualizes public value as a dynamic process of (dis)aggregation, where multiple and conflicting value regimes coexist and are negotiated through evaluative practices. This allows us to explain not only the multiplicity of value generated by such organizations, but also the tensions, asymmetries, and exclusions through which that value comes into being.

The paper is based on an exploratory qualitative case study of Kebula S.r.l., an academic spin-off operating in data engineering and AI. Adopting a valuographic approach (Dussauge et al., 2015), the analysis examines how value is constructed, justified, and governed through everyday practices. Findings show that AI systems act as infrastructures of valuation that reorganize what counts as legitimate impact, redistributing visibility across economic, social, and knowledge domains. Public value emerges as a contested and relational accomplishment rather than a stable outcome. The paper contributes to knowledge management, public value theory, and AI-entrepreneurship literature by showing that intelligent transformation shapes not only how value is created, but also how it is made visible, accountable, and legitimate. Empirically, the case highlights how digital knowledge systems mediate innovation, learning, failure, and infrastructural flexibility in a hybrid university-based organization over time and across organizational contexts.

Keywords: Academic spin-offs; Artificial Intelligence; Public value; Valuography; Knowledge ecosystems

Paper type: Academic Research Paper

1 Introduction

Artificial intelligence (AI) is increasingly celebrated as a strategic lever of digital transformation. Across sectors, AI promises to improve decision-making, accelerate innovation, enhance prediction, and unlock new forms of organizational intelligence (Agarwal et al., 2021; Shepherd and Majchrzak, 2022; Fossen et al., 2024). Yet, the expansion of AI also raises a deeper question that cannot be reduced to efficiency or competitiveness alone: what kinds of value are being produced through AI-mediated knowledge processes, for whom, and through which evaluative infrastructures? This question is especially pressing for organizations that operate at the intersection of economic, scientific, and societal mandates, such as academic spin-offs (Iacobucci and Micozzi, 2015; Miranda et al., 2018).

Academic spin-offs occupy a distinctive position in contemporary knowledge economies. They translate scientific research into entrepreneurial initiatives, but they also remain tied to broader public missions associated with universities, including education, research, and territorial development (Abramo et al., 2012; Li et al., 2022). They are therefore not simply firms; they are hybrid organizational forms embedded in knowledge ecosystems where universities, businesses, students, technologies, investors, and public institutions interact (Battilana and Lee, 2014; Mele et al., 2023). In such settings, the creation of value cannot be reduced to profit, growth, or market share. It also concerns the circulation of expertise, the formation of talent, the enhancement of teaching and research, the reshaping of local capabilities, and the production of organizational and societal futures (Davenport and Prusak, 1998; Zack, 1999).

This paper argues that AI intensifies and complicates these dynamics. It does so not merely by augmenting entrepreneurial decision-making, but by actively reconfiguring the knowledge practices through which value is enacted, made visible, and judged (Shepherd and Majchrzak, 2022; Fossen et al., 2024). AI does not simply process data. It participates in the construction of what counts as relevant knowledge, meaningful performance, legitimate impact, and actionable opportunity (Miller, 1998; Mouritsen et al., 2009). This is why it is insufficient to speak of AI as a neutral enabler of innovation. Instead, AI must be understood as part of the evaluative

infrastructures through which organizations know, decide, justify, and govern (Kornberger et al., 2015; Dussauge et al., 2015).

To capture these dynamics, the paper draws on public value theory and the sociology of valuation. Public value theory has long emphasized that organizations, especially those connected to public institutions, generate value beyond economic returns (Moore, 1995; 2013; Bracci et al., 2019). However, the concept of public value remains ambiguous and often overly normative (Steccolini, 2019). The sociology of valuation helps unpack this ambiguity by shifting the analytical focus from value as an object to valuing as a situated practice (Stark, 2009; Dussauge et al., 2015). Combining these perspectives, the paper conceptualizes public value as a process of (dis)aggregation. Value is continuously assembled through narratives, metrics, classifications, and infrastructures, but it is also disassembled through exclusions, failures, trade-offs, and forms of invisibility (Miller, 1998; Carruthers and Kim, 2011; Kornberger et al., 2015). What is omitted or backgrounded is not external to value creation; it is constitutive of it.

This argument is developed through an exploratory case study of Kebula S.r.l., an academic spin-off of the University of Salerno specializing in data engineering, analytics, and artificial intelligence. Kebula is a particularly revealing case because it sits at the intersection of university-based knowledge production and market-facing AI solutions. Its activities range from data platforms and analytics to audience measurement, recommendation systems, and process automation. In this context, AI is not only a technical capability but a mode of organizing, a way of making knowledge travel, and a mechanism through which different forms of value are articulated and contested.

The paper shows that AI-driven academic spin-offs generate public value not because they harmoniously combine economic and social goals, but because they operate as sites where multiple value regimes are continuously negotiated. In doing so, they reveal both the promise and the tensions of intelligent transformation in hybrid knowledge organizations.

2 Theoretical background

2.1 Public value beyond stable outcomes

Public value theory emerged as an attempt to rethink the role of public organizations and public managers beyond narrow efficiency criteria (Moore, 1995; 2013). Rather than focusing exclusively on compliance, output measures, or incentive alignment, the notion of public value emphasizes the broader benefits that organizational action can generate for society (Bracci et al., 2019; Steccolini, 2019). Yet, despite its appeal, public value remains an unstable and often elusive construct. It is difficult to define, difficult to operationalize, and difficult to separate from the evaluative practices through which it is claimed and recognized.

This instability matters. Public value is often invoked as if it were self-evident: a desirable societal outcome that organizations can pursue, measure, and report. However, such a view risks reifying value into a fixed object. As a growing body of work in public sector accounting and management suggests, public value is not simply "there" waiting to be captured. It is actively constituted through managerial, political, and calculative practices (Miller, 1998; Bracci et al., 2019; Steccolini, 2019). Budgets, indicators, rankings, targets, reports, and impact frameworks do not merely describe value; they participate in its production (Mouritsen et al., 2009; Carruthers and Kim, 2011).

From this perspective, public value is not only plural but also contested. It involves multiple stakeholders with different evaluative principles. What counts as valuable for a university manager, a public policymaker, a spin-off founder, a student, or a local ecosystem actor may not coincide. This multiplicity makes public value less a coherent outcome than an arena of negotiations, compromises, and tensions.

2.2 Valuation and the (dis)aggregation of value

The sociology of valuation provides conceptual tools to address this complexity. Rather than assuming that value is an intrinsic property of things or actions, valuation studies examine the practices through which value is enacted (Stark, 2009; Dussauge et al., 2015; Kornberger et al., 2015). This includes the devices, classifications, tests, attachments, and infrastructures that make things count (Dussauge et al., 2015). The central analytical move is to shift from asking "what is value?" to asking "how does valuing happen?"

This move is especially useful for understanding public value in hybrid contexts such as academic spin-offs. These organizations are subject to multiple evaluative logics at once. They are judged as businesses, as knowledge transfer vehicles, as educational spaces, as territorial development actors, and as embodiments of the university's third mission. Each of these logics carries its own criteria, temporalities, and forms of visibility.

To capture this multiplicity, this paper mobilizes the concept of (dis)aggregation. Value is aggregated when elements are connected, grouped, and represented as contributing to a broader social good. But value is also disaggregated when some elements are left out, purified, deferred, or rendered invisible. (Dis)aggregation does not simply mean fragmentation. It refers to the constitutive work of separation through which value formation becomes possible. In other words, every value claim is sustained by a corresponding background of exclusions, silences, and ambiguities.

This is why additive models of value, such as some versions of blended value thinking, are only partially helpful. They correctly recognize that economic and social values are connected, but they often assume that these values can be harmoniously combined (Emerson, 2003; Nicholls, 2009). The valuographic perspective adopted here is more skeptical. It assumes that value regimes may nourish one another, but also clash, overflow, and destabilize each other (Stark, 2009; Kornberger et al., 2015).

2.3 Knowledge management beyond resource metaphors

Knowledge management literature has long stressed the importance of knowledge as a strategic asset (Davenport and Prusak, 1998; Zack, 1999). Classic contributions frame knowledge as something that can be captured, codified, transferred, and leveraged for competitive advantage (Davenport and Prusak, 1998). This literature has been highly influential in highlighting the organizational centrality of learning, expertise, and intellectual capital. Yet, it has also often relied on a relatively stable image of knowledge as a resource that exists independently of the practices through which it is recognized and mobilized.

More practice-oriented perspectives complicate this picture by suggesting that knowledge is not simply stored and moved but enacted in situated activities (Mouritsen et al., 2009; Kornberger et al., 2015). What counts as valid knowledge depends on organizational routines, interpretive frames, technologies, and material infrastructures (Dussauge et al., 2015). Knowledge is thus inseparable from organizational forms of valuation. To know is also to rank, classify, prioritize, and render some signals more consequential than others.

This point becomes crucial in AI-intensive settings. AI systems are not external tools acting upon stable pools of knowledge. They participate in the production of knowledge by selecting, cleaning, weighting, predicting, and recommending (Shepherd and Majchrzak, 2022; Fossen et al., 2024). Accordingly, intelligent knowledge systems must be understood not only as knowledge infrastructures but also as evaluative infrastructures (Mele et al., 2023).

2.4 AI and entrepreneurship beyond augmentation

Emerging research on AI and entrepreneurship highlights how AI can assist opportunity recognition, market analysis, experimentation, and business model development (Shepherd and Majchrzak, 2022; Fossen et al., 2024). Entrepreneurs can use AI to process more information, simulate scenarios, identify patterns, and accelerate strategic choices (Fossen et al., 2024). This literature is important because it moves beyond purely technical accounts of AI and shows how AI affects entrepreneurial cognition and organizing.

However, much of this literature still treats AI primarily as an enhancer of pre-existing entrepreneurial processes. The underlying assumption is that entrepreneurship remains fundamentally human, while AI improves speed, accuracy, and scope. Such an approach risks underestimating the performative role of AI.

AI systems do not simply help entrepreneurs discover opportunities; they shape what becomes intelligible as an opportunity in the first place (Shepherd and Majchrzak, 2022). They do not only support decision-making; they configure the informational environments within which decisions become thinkable and defensible (Miller, 1998; Mouritsen et al., 2009). They do not merely optimize action; they participate in defining what should be optimized.

This is where insights from performativity and calculative practices become useful. If calculative devices can act at a distance by rendering distant phenomena knowable and governable, then AI can be seen as a particularly powerful extension of such capacity. Yet, unlike conventional accounting devices, AI introduces forms of prediction, automation, and opacity that intensify both the promises and the ambiguities of organizational valuation.

3 Methodology

3.1 Research design

This study adopts an exploratory qualitative case study design (Yin, 2018). A case study is appropriate because the phenomenon under investigation is complex, context-dependent, and difficult to isolate from the

organizational and institutional settings in which it unfolds (Yin, 2018). The aim is not statistical generalization, but analytical development: to generate theoretically informed insights into how AI-driven academic spin-offs enact public value in practice.

The selected case is Kebula S.r.l., an academic spin-off of the University of Salerno. Academic spin-offs are particularly relevant for examining knowledge transfer, innovation, and local development because they operate at the boundary between academic research and entrepreneurial activity (Iacobucci and Micozzi, 2015; Miranda et al., 2018). Kebula is registered as an innovative start-up and operates in the fields of data engineering, analytics, and AI. The case was chosen because it is strongly embedded in the university's entrepreneurial ecosystem while simultaneously operating in highly data-intensive and market-facing sectors, including media and broadcasting. This makes it particularly suitable for exploring how AI-mediated knowledge practices intersect with public value creation.

The data were analysed through a valuographic lens (Dussauge et al., 2015). Valuography does not assume stable categories of value in advance. Instead, it follows how actors describe, justify, prioritize, and contest what matters (Dussauge et al., 2015; Stark, 2009).

3.2 Empirical context

The case emerges from the local environment of the University of Salerno and from the role of intermediary structures supporting technology transfer and spin-off development. Like many university-based entrepreneurial initiatives, Kebula sits in a space where academic research, entrepreneurship, and institutional support intersect. It is therefore not simply a business venture but part of a broader ecology of innovation, training, and public investment.

Kebula's mission is to help clients "datafy" their businesses. This means transforming unprocessed data into structured, analyzable, and strategically useful information. Its work includes the development of end-to-end data platforms, AI-enabled analytics, machine learning models, automation solutions, and content-related tools. In the broadcasting sector, for instance, the company has worked on audience analytics and data infrastructures for major players such as RTL 102.5.

3.3 Data collection and analysis

The analysis draws on semi-structured interviews with founders and key organizational actors engaged in the spin-off's activities. These interviews focused on the origin of the spin-off, its technological practices, its relationship with the university, and the various forms of value perceived to be generated through its work. Interview material was complemented by contextual analysis of the company's activities, its positioning, and its project descriptions.

The data were analysed through a valuographic lens (Dussauge et al., 2015; Stark, 2009), which follows how actors describe, justify, prioritize, and contest what matters.

The analytical process proceeded iteratively. First, interview material was read to identify explicit and implicit references to value, impact, and public significance. Second, these references were grouped into recurring themes, including knowledge transfer, innovation, talent formation, failure, educational payback, research enhancement, infrastructural flexibility, and bureaucratic tension. Third, these themes were interpreted as interrelated value regimes rather than isolated outcomes. Finally, the analysis examined how AI and data practices mediated the enactment of these regimes.

4 Findings

4.1 Innovation as knowledge translation and public contribution

A first dimension of public value concerns the translation of academic knowledge into organizational practice. Interviewees repeatedly stressed that the spin-off was not engaged in "business as usual." This expression signalled more than a start-up ethos. It pointed to a form of entrepreneurship rooted in experimentation, co-engineering, and the mobilization of specialized knowledge for broader organizational and societal purposes.

The CTO articulated this clearly:

"The most immediate public value we contribute to is the transfer of experiences and expertise, an entire technical and cultural baggage to the productive world. The aim is to innovate, create new opportunities. This is also what is considered as the substantial value justifying the university's third mission."

What is striking in this statement is that innovation is not framed narrowly as product development or firm performance. It is framed as transfer, opportunity creation, and contribution to the university's public mission. The value of the spin-off resides not only in its capacity to commercialize knowledge but in its role as a conduit through which academic expertise acquires broader relevance.

This suggests that innovation itself is being valued in a public register. The spin-off's work is understood as contributing to a larger ecosystem of organizational learning and capability building (Zack, 1999; Mele et al., 2023). The public value claim therefore does not emerge after the fact through ex post measurement; it is embedded in the very narrative through which the organization makes sense of its existence.

4.2 Datafication, AI, and the construction of actionable reality

A second key finding concerns the role of datafication and AI as practical means through which value is enacted. Kebula's core activity is not simply to analyse data, but to construct datafied organizational realities. This involves collecting heterogeneous sources of information, structuring them into meaningful categories, and rendering them available for decision-making.

The company's activities span big data analysis, machine learning, process automation, and content-related applications. In the broadcasting sector, for example, Kebula's work includes audience analysis, content recommendation, and the development of data platforms capable of integrating signals from production, contribution, packaging, distribution, and consumption.

The spin-off developed a data platform intended not merely to gather data, but to create a new model of radio audience measurement. This is analytically important because audience measurement is not a neutral task. It does not simply reveal what listeners do. It defines what counts as an audience, which interactions matter, how engagement is represented, and what signals become strategically consequential. In other words, AI and data platforms do not merely describe organizational reality; they participate in formatting it.

4.3 Failure as ecosystem learning and redistributed value

A particularly interesting finding concerns the revaluation of failure. In conventional entrepreneurial discourse, failure is often treated as value destruction. By contrast, interviewees suggested that failures in spin-off activity can still generate public value when viewed from a systemic perspective.

The CTO observed:

"Increasing the probability of success in financing a spin-off initiative is undoubtedly an advantage because it allows more people to approach these challenges. Even if spin-offs were to fail, they wouldn't actually be failures because such failures would be absorbed by the system. Providing opportunities for Italian talents to express themselves in an environment close to the market through a spin-off… represents comprehensive value because those individuals, even if the spin-off were to fail, remain within the ecosystem and become a valuable resource."

This passage is theoretically significant. It shows that value is being redistributed away from the isolated unit of the firm toward the broader ecosystem. Failure is no longer evaluated solely at the level of business performance. It is reinterpreted as a mechanism through which learning, exposure, and capability building are retained within the environment.

This challenges linear and firm-centric models of entrepreneurial evaluation. Public value here emerges precisely through disaggregation: what appears as failure at one level may become productive at another. This also suggests that conventional metrics such as turnover, cash flow, or profit may be insufficient for evaluating the public significance of AI-driven academic entrepreneurship (Moore, 2013; Bracci et al., 2019).

4.4 Anti-bureaucratic flexibility and the politics of infrastructure

A final major finding concerns the role of the spin-off as an anti-bureaucratic mechanism. Interviewees emphasized that the organizational form of the spin-off allowed greater flexibility than a university department would permit, especially in relation to computing infrastructure and cloud-based resources.

The CTO put it bluntly:

"If today I wanted, within the Italian university, to purchase a cloud service from Google or Amazon, this would be practically impossible because, normatively, it's considered public expenditure… But once I have this machine, if I don't use it at full capacity, I'm destroying value… less structure, less bureaucracy, and more action help create value!"

This statement makes visible a central tension in AI-intensive public value creation. Governance arrangements designed to ensure accountability can inhibit the very flexibility needed to deploy AI effectively.

The spin-off form becomes a workaround: an organizational infrastructure that absorbs friction and enables experimentation (Battilana and Lee, 2014; Miranda et al., 2018).

This finding reveals that value creation beyond profit depends not only on entrepreneurial intent or technological capability, but also on the material and regulatory infrastructures that make action possible. Public value is therefore intertwined with the politics of organizational form.

5 Discussion

5.1 AI as infrastructure of valuation

The findings show that AI-driven academic spin-offs should not be understood merely as knowledge transfer vehicles or innovation intermediaries. They are sites where value is continuously made, tested, redistributed, and contested. This process is deeply mediated by AI and data infrastructures.

AI acts here as an infrastructure of valuation in at least three senses. First, it structures visibility. Through data collection, model building, dashboards, and analytics, AI foregrounds certain dimensions of organizational life and renders them actionable. Listener behaviour, process efficiency, engagement signals, predictive patterns, and usage dynamics become visible in standardized formats. Yet, what AI makes visible is never neutral. It depends on prior choices about what to measure, how to measure it, and what to optimize.

Second, AI extends organizational action at a distance. Once translated into data and models, audiences, processes, and organizational states can be monitored and acted upon remotely. This echoes classic analyses of calculative infrastructures, but AI intensifies these capacities by combining scale, speed, prediction, and automation. In the Kebula case, action at a distance is not abstract. It is materialized in audience analytics, content recommendation, and platform-based representations that travel across organizational boundaries.

Third, AI produces a paradoxical combination of transparency and opacity. On the one hand, metrics and dashboards increase traceability and comparative visibility. On the other hand, the very complexity of AI systems can obscure how specific outputs are generated. Responsibility becomes more difficult to assign, not less. Accountability is redistributed across data engineers, algorithms, clients, and infrastructures.

5.2 Public value as (dis)aggregation rather than harmony

These findings support a reconceptualization of public value as a process of (dis)aggregation. Public value is not a stable stock of benefits produced by the spin-off and then measured by external observers. It is a precarious accomplishment constituted through the alignment, tension, and partial reconciliation of multiple value regimes.

This is where the paper departs from additive or harmonizing accounts. Economic value, social value, educational value, research value, and infrastructural value do not simply accumulate. They interact in uneven ways. Sometimes they reinforce one another, as when industrial engagement improves teaching and research. Sometimes they conflict, as when performance-oriented data systems marginalize less measurable dimensions such as editorial judgment or cultural plurality. Sometimes they are redistributed across levels, as when apparent firm-level failure generates ecosystem-level learning.

The concept of (dis)aggregation allows us to see these movements without forcing them into a unified metric. It also helps explain why public value is so difficult to stabilize conceptually: because it is generated precisely through the coexistence of complementarity and contradiction.

5.3 Contributions to knowledge management

The paper contributes to knowledge management by arguing that AI-mediated knowledge processes are inseparable from valuation processes. Traditional knowledge management often emphasizes capture, sharing, transfer, and leverage. These remain important concerns, but the present analysis shows that what counts as knowledge in AI-driven settings depends on evaluative infrastructures.

In Kebula's case, datafication is not only about extracting knowledge from organizational reality. It is about formatting reality into knowledge that can circulate, be optimized, and justify action. Data platforms, model outputs, and KPIs are not simply knowledge containers. They are valuation devices that rank some forms of knowing over others.

This implies a more critical view of intelligent knowledge systems. Their significance lies not only in improving knowledge flows but in reorganizing the criteria through which knowledge becomes relevant. Knowledge ecosystems, in turn, are not merely networks of exchange. They are arenas where epistemic and evaluative orders are continuously produced together.

5.4 Contributions to public value theory

The paper contributes to public value theory by moving away from a normative and outcome-focused understanding of public value toward a relational and processual one. Rather than asking whether an organization creates public value in the abstract, the analysis asks how value is enacted through concrete practices, and which forms of visibility, exclusion, and redistribution sustain those claims.

This move is important because it prevents public value from becoming an empty managerial slogan. If public value is treated as a self-evident organizational aspiration, it becomes difficult to analyse critically. By contrast, treating public value as (dis)aggregation makes it possible to study how claims of public benefit are assembled, how they travel, and what they leave aside.

The paper also shows that public value in hybrid organizations is not merely a matter of balancing market and mission. It is an effect of socio-technical organizing. AI infrastructures, cloud services, educational feedback loops, talent retention mechanisms, and entrepreneurial narratives all participate in making public value plausible and durable.

5.5 Contributions to AI-entrepreneurship literature

The AI-entrepreneurship literature has rightly emphasized the promise of AI for opportunity recognition and venture development. This paper adds a critical dimension by showing that AI does more than enhance entrepreneurial cognition. It actively configures the evaluative environments within which entrepreneurship unfolds.

In this sense, AI is not just an aid to entrepreneurship but part of the organizational politics of value. It influences which opportunities are seen, which outcomes are prioritized, and which forms of legitimacy are available. This is particularly consequential in academic entrepreneurship, where ventures are accountable not only to markets but also to universities, students, and public narratives of innovation.

6 Conclusion

This paper has examined how an AI-driven academic spin-off enacts public value within a knowledge ecosystem. Drawing on public value theory, valuation studies, knowledge management, and AI-entrepreneurship literature, it has argued that public value is best understood not as a stable or additive outcome but as a process of (dis)aggregation.

The paper makes three main contributions. First, it contributes to knowledge management by showing that intelligent knowledge systems are inseparable from valuation processes. Second, it contributes to public value theory by conceptualizing public value as a relational and contested accomplishment. Third, it contributes to AI-entrepreneurship research by showing that AI actively shapes the evaluative infrastructures through which entrepreneurship becomes meaningful and legitimate.

7 Ethics declaration

No ethical approval was required.

8 AI declaration

AI tools were used to support English language refinement. All intellectual content and analysis remain the responsibility of the author.

References

Abramo, G., D'Angelo, C.A. and Ferretti, M. (2012) 'Academic entrepreneurship and research productivity', R&D Management, 42(3), pp. 1–15.

Agarwal, R., Guodong, G., DesRoches, C., Jha, A.K. and Gans, D. (2021) 'The digital transformation of healthcare: Current status and the road ahead', Information Systems Research, 32(2), pp. 1–18.

Battilana, J. and Lee, M. (2014) 'Advancing research on hybrid organizing: Insights from the study of social enterprises', Academy of Management Annals, 8(1), pp. 397–441.

Bracci, E., Papi, L., Bigoni, M., Deidda Gagliardo, E. and Bruns, H. (2019) 'Public value and public sector accounting research: A structured literature review', Journal of Public Budgeting, Accounting & Financial Management, 31(1), pp. 103–136.

Carruthers, B. G., & Kim, J.-C. (2011). The sociology of finance. Annual Review of Sociology, 37, 239–259.

Davenport, T. and Prusak, L. (1998) Working Knowledge. Boston, MA: Harvard Business School Press.

Dussauge, I., Helgesson, C.F. and Lee, F. (2015) Value Practices in the Life Sciences and Medicine. Oxford: Oxford University Press.

Emerson, J. (2003) 'The blended value proposition: Integrating social and financial returns', California Management Review, 45(4), pp. 35–51.
Fossen, F.M., McLemore, T. and Sorgner, A. (2024) 'Artificial intelligence and entrepreneurship', Foundations and Trends in Entrepreneurship, 20(8), pp. 781–904.
Iacobucci, D. and Micozzi, A. (2015) 'How to evaluate the impact of academic spin-offs on local development: An empirical analysis of the Italian case', The Journal of Technology Transfer, 40(3), pp. 434–452.
Kornberger, M., Justesen, L. and Mouritsen, J. (2015) Making Things Valuable. Oxford: Oxford University Press.
Li, H., Yang, X. and Cai, X. (2022) 'Academic spin-off activities and research performance: The mediating role of research collaboration', The Journal of Technology Transfer, 47, pp. 1037–1069.
Mele, C., Russo-Spena, T. and Kaartemo, V. (2023) 'Value creation in digital service ecosystems', Journal of Business Research, 155, 113372.
Miller, P. (1998) 'The margins of accounting', European Accounting Review, 7(4), pp. 605–621.
Miranda, F.J., Chamorro, A. and Rubio, S. (2018) 'Re-thinking university spin-off: A critical literature review and a research agenda', The Journal of Technology Transfer, 43(4), pp. 1007–1038.
Moore, M.H. (1995) Creating Public Value. Cambridge, MA: Harvard University Press.
Moore, M.H. (2013) Recognizing Public Value. Cambridge, MA: Harvard University Press.
Mouritsen, J., Hansen, A. and Hansen, C.Ø. (2009) 'Short and long translations: Management accounting calculations and innovation management', Accounting, Organizations and Society, 34(6), pp. 738–754.
Nicholls, A. (2009) '"We do good things, don't we?": "Blended value accounting" in social entrepreneurship', Accounting, Organizations and Society, 34, pp. 755–769.
Shepherd, D.A. and Majchrzak, A. (2022) 'Machines augmenting entrepreneurs: Opportunities (and threats) at the nexus of artificial intelligence and entrepreneurship', Journal of Business Venturing, 37(4), 106227.
Stark, D. (2009) The Sense of Dissonance. Princeton, NJ: Princeton University Press.
Steccolini, I. (2019) 'Accounting and the post-new public management: Re-considering publicness in accounting research', Accounting, Auditing & Accountability Journal, 32(1), pp. 255–279.
Yin, R.K. (2018) Case Study Research and Applications. Thousand Oaks, CA: Sage.
Zack, M.H. (1999) 'Developing a knowledge strategy', California Management Review, 41(3), pp. 125–145.

Towards an Integrated Digital Twin Maturity Assessment Framework for Coastal Cultural Heritage Resilience

Sharmila Gope[1], Claudio Pascarelli[1], Mattia Mangia[1], Giovanni Coppini[2], Ivan Federico[2], Salvatore Causio[2], Marco Boetti[2], Federica Stella Blasi[2]

[1] University of Salento, Lecce, Italy

[2] CMCC Foundation - Euro-Mediterranean Center on Climate Change, Italy

Abstract

Coastal and underwater cultural heritage is increasingly exposed to climate-driven stressors, including sea-level rise, shoreline erosion, salinization, extreme weather events, and marine environmental change. These pressures are accelerating the degradation of vulnerable archaeological sites and historic coastal assets, while also exposing the limits of conventional documentation and conservation approaches based on periodic surveys and fragmented datasets. In this context, Digital Twins (DTs) are emerging as a promising paradigm for heritage monitoring and preservation, as they can integrate digital representation, sensor-based observation, environmental information, and predictive analysis within a dynamic and potentially interoperable system. However, current applications in cultural heritage remain uneven and frequently technology-centred, with limited attention to institutional embedding, environmental coupling, and long-term knowledge continuity. To address this gap, this study proposes a multidimensional maturity assessment framework for Digital Twins in coastal and underwater cultural heritage. The research adopts a sequential mixed-method design. First, a bibliometric analysis of 346 publications published between 2010 and 2025 was conducted using Scopus and Bibliometrix in RStudio to identify thematic trends, research fragmentation, and underrepresented dimensions in the literature. The results show a rapidly growing field, but still organized around weakly connected clusters, especially between digital representation, environmental modelling, and governance- or knowledge-related themes. Second, these findings were translated into a five-dimensional maturity framework comprising Technical, Structural, Environmental, Governmental, and Social dimensions. The proposed framework redefines DT maturity as a multidimensional preservation condition rather than a purely technological stage. It enables the assessment of whether a DT can function not only as a digital model, but also as an interoperable, governable, and sustainable preservation system. The findings indicate that current initiatives are advancing more strongly in technical and structural terms, while environmental, governmental, and social dimensions remain significantly underdeveloped. The framework, therefore, offers an evidence-based tool for identifying implementation gaps, supporting strategic planning, and guiding the more balanced development of Digital Twins for long-term heritage preservation.

Keywords: Digital Twins, Coastal Cultural Heritage, Risk Management, Maturity Assessment, Knowledge Sharing

Paper type: Academic Research Paper

1 Introduction

Underwater and coastal cultural heritage (U&C-CH) (including shipwrecks, submerged settlements, harbours, and maritime fortifications) constitutes an integral part of the cultural heritage of humanity and preserves unique evidence of past societies, maritime connections, and long-term human–environment interactions. As recognized by the 2001 UNESCO Convention (UNESCO, 2001), these evidences from the past are not only archaeological resources but also shared cultural assets whose preservation should be prioritized in situ whenever possible. Today, however, U&C-CH is increasingly exposed to climate-driven hazards, including sea-level rise, coastal flooding, erosion, storm intensification, ocean warming, and acidification, all of which can accelerate material decay, destabilize seabeds, and increase the physical vulnerability of both coastal and submerged sites. These pressures are particularly urgent in the Mediterranean, where coastal heritage is densely concentrated and where exposure to flooding and erosion is projected to increase substantially during the twenty-first century (Gregory et al., 2022; Reimann et al., 2018; UNESCO, 2001). Conventional conservation approaches, typically based on periodic surveys, isolated documentation campaigns, and reactive interventions, are increasingly inadequate for addressing hazards that are cumulative, dynamic, and spatially interconnected. Recent reviews show that research on climate change and cultural heritage has grown rapidly, yet the field remains fragmented across disciplines, institutions, and regions, and only a limited proportion of studies is effectively translated into adaptation planning, management protocols, or long-term monitoring strategies. For underwater heritage in particular, the challenge is amplified by site inaccessibility, the scarcity of baseline datasets, and the difficulty of integrating environmental observations with archaeological documentation to

support timely decision-making and risk prioritization (Gregory et al., 2022; Orr et al., 2021). Digital Twins (DTs) are emerging as a relevant framework for monitoring and managing heritage assets, as they integrate geometric and semantic data, sensor inputs, environmental information, and simulation models within a continuously updated digital representation. In cultural heritage, the transition from HBIM (Heritage-Building Information Modeling) to DTs marks a shift from static documentation to dynamic systems that support condition assessment, predictive maintenance, and scenario analysis, enabling more timely and informed conservation decisions, particularly in rapidly changing environments (Grazianova & Mesaros, 2024; Hou et al., 2024). However, current applications remain uneven and generally immature. Key challenges include managing heterogeneous data, structuring semantic information, integrating enabling technologies, and embedding DTs into operational workflows. Recent studies also highlight that maturity depends not only on technical capabilities, but on interoperability, institutional adoption, and effective knowledge exchange (Hou et al., 2024; Klar et al., 2024). To address this gap, this research develops a multidimensional maturity assessment framework for Digital Twins in coastal resilience and U&C-CH. Rather than evaluating DTs development only through technical deployment, the framework considers the interdependence of technical, structural, environmental, governmental, and social dimensions. Grounded in a bibliometric review of 346 publications, the framework is intended to support archaeologists, engineers, heritage managers, and policymakers in assessing not only whether a DT can be built, but whether it can function as an interoperable, governable, and sustainable system for long-term heritage preservation.

2 Background

2.1 Digital twins: definition, architecture, and heritage applications

Digital Twins (DTs) are commonly understood not as generic 3D models, but as systems in which a physical asset and its digital counterpart are connected through automated data exchange. In the literature, a DT is commonly distinguished from a simple digital model, which has no automated link with the physical asset, and from a digital shadow, in which data flow is unidirectional. A DT, by contrast, implies at least the possibility of continuous or recurrent synchronization between physical and digital states. Although DT architectures vary across domains, most definitions converge on a recurring set of elements: a physical asset, a digital representation, data acquisition systems, data storage and processing layers, communication infrastructure, and analytical or simulation services supporting visualization, monitoring, and prediction (Fuller et al., 2020). In cultural heritage, this distinction is especially important because many systems described as “digital twins” are, in practice, advanced documentation environments rather than fully synchronized twins. Current research consistently indicates that HBIM can provide a crucial basis for DT development by structuring geometric, historical, and semantic information, but that HBIM alone should not be equated with a DT. What distinguishes a heritage DT is the integration of documentation with monitoring, updating, analysis, and decision support, so that the digital environment can assist conservation practice rather than merely record the asset. In this sense, DTs extend heritage documentation toward condition assessment, predictive maintenance, and operational management (Grazianova & Mesaros, 2024; Hou et al., 2024). For cultural heritage applications, the real added value of DTs lies less in visual replication alone than in the capacity to integrate heterogeneous information (e.g. survey outputs, monitoring records, diagnostics, conservation data, and interpretive knowledge), within a coherent and interoperable information environment. This issue has become central in recent work on the Heritage Digital Twin ontology, which proposes a CIDOC CRM–based (CIDOC Conceptual Reference Model) semantic framework capable of linking diverse documentation systems while preserving interoperability across institutions and research domains (Niccolucci et al., 2023). Such an approach can become particularly relevant for underwater and coastal cultural heritage, where archaeological evidence, environmental observations, and conservation actions are often fragmented across incompatible datasets, tools, and organizations.

2.2 Existing maturity assessment frameworks: capabilities and limitations

A range of maturity assessment frameworks can inform the evaluation of Digital Twin development, but each addresses only a partial dimension of the problem. Technology Readiness Levels (TRLs) remain useful for measuring the progression of technology from early research to operational deployment (European Commission. Directorate General for Research and Innovation, 2017). The Capability Maturity Model Integration (CMMI) focuses instead on organizational capability and process improvement (Chrissis et al., 2003) while BIM maturity models or matrix evaluate the quality, repeatability, and predictability of information modelling practices (Underwood & Isikdag, 2010). These frameworks are therefore valuable as reference points, but they

were not originally designed to assess the combined technical, semantic, environmental, and organizational demands of heritage-oriented DT systems. This limitation becomes clearer in heritage applications. Reviews of DT adoption in heritage facilities management describe the field as still preliminary and identify recurring implementation bottlenecks (Grazianova & Mesaros, 2024; Hou et al., 2024): the need to manage very large and heterogeneous datasets, the difficulty of building structured semantic databases, the integration of HBIM with IoT, etc. In other words, technical deployment alone does not provide a reliable indicator of maturity. A system may include advanced modelling and sensing components, yet remain weak in interoperability, documentation structure, routine usability, or institutional uptake (Hou et al., 2024; Niccolucci et al., 2023). More broadly, the aforementioned studies' maturity argues that interoperability should be treated as a defining feature of advanced digital twins rather than as a secondary add-on. This observation is highly relevant for coastal and underwater heritage, where effective preservation depends on the coordinated use of environmental monitoring, spatial analysis, semantic documentation, and multi-actor management. At the same time, climate-change and heritage research have repeatedly shown that fragmentation across disciplines, institutions, and adaptation practices remains a major barrier to effective action. For this reason, a maturity model for heritage DTs cannot be limited to technical readiness or geometric richness alone, but should also consider structural coherence, environmental coupling, organizational preparedness, and knowledge sharing across stakeholders (Klar et al., 2024).

2.3 Research gap: multidimensional assessment framework for heritage digital twins

In the context of cultural heritage, Digital Twins should operate not only as technical systems but also as long-term infrastructures (Barzaghi et al., 2024; Gabellone, 2022) for documentation, monitoring, interpretation, and decision support across heterogeneous actors and datasets. A major gap concerns interoperability and information continuity. Recent work on FAIR-by-design (Findibale, Accesible, Interoperable, Reusable) digital twins in cultural heritage has shown that long-term usability depends on much more than 3D acquisition quality: it also requires pre-structured workflows, provenance tracking, open standards, and metadata models capable of supporting reuse across platforms and institutions (Barzaghi et al., 2024). Likewise, emerging ontology-based approaches for cultural heritage preservation emphasize that static documentation models are insufficient when the objective is predictive monitoring and preventive conservation. Frameworks such as PANOPTES have therefore been developed to integrate sensor observations, diagnoses, risk predictions, and conservation decisions within a common semantic environment (Pavlidis et al., 2025). These contributions make clear that maturity in heritage DTs cannot be reduced to modelling sophistication alone, but must include the ability to sustain interoperable, traceable, and operational data ecosystems.

Against this background, the present study addresses the following research questions:

- *RQ1. Which dimensions are necessary to assess the maturity of Digital Twins for coastal and underwater cultural heritage beyond conventional measures of technological readiness?*
- *RQ2. How can a multidimensional maturity framework support heritage practitioners and decision-makers in identifying implementation gaps, comparing initiatives, and guiding the sustainable deployment of Digital Twins in real-world preservation contexts?*

These research questions directly motivate the DTs' five dimensions proposed in this study. The Technical dimension addresses data acquisition, sensing, updating workflows, simulation, and risk-processing capabilities. The Structural dimension evaluates the quality of geometric representation, semantic enrichment, provenance, and interoperability of the digital asset. The Environmental dimension assesses the extent to which a Digital Twin incorporates dynamic coastal or underwater observations, hazard indicators, and predictive environmental modelling. The Governmental dimension considers governance arrangements, stakeholder coordination, organizational readiness, and the integration of a Digital Twin into decision-making processes. Finally, the Social dimension captures documentation quality, training, knowledge transfer, and the capacity of institutions and practitioners to maintain and reuse the system over time.

3 Research Methodology

This study adopts a sequential mixed-method design to develop a multidimensional maturity assessment framework for Digital Twins (DTs) in coastal and underwater cultural heritage. The research was therefore structured in two consecutive phases: (i) a bibliometric analysis to identify thematic patterns and implementation gaps in the literature, and (ii) the development of a multidimensional framework grounded in those findings. Additionally, a weighted scoring method is proposed to compute the overall maturity score.

3.1 Data source and search strategy

The bibliometric dataset was compiled from Scopus in February 2026, which has been selected for its multidisciplinary coverage and suitability for reproducible bibliometric analysis (Mongeon & Paul-Hus, 2016). The search strategy combined terms related to Digital Twins, maturity, BIM/HBIM/GIS, cultural heritage, and related monitoring and preservation domains, employing the following query after several iterations aimed to properly address sources related with the field of investigation: TITLE-ABS-KEY(("digital twin*" or "digital shadow" or "cyber-physical system*") AND ("maturity" or "readiness" or "capabilit*") AND ("BIM" or "HBIM" or "GIS" or "BIM/GIS" OR "cultural heritage" OR "climate change" or "sea level")). The time window covered publications from 2010 to 2025. Following a screening of records to exclude incorrect data (e.g., Conference Review records), the final dataset included 346 publications. The records were analysed in Rstudio using the Bibliometrix package (Aria & Cuccurullo, 2017). The bibliometric phase served two purposes. First, it mapped the growth and thematic configuration of the field through annual scientific production and keyword co-occurrence analysis. Second, and more importantly, it was used to detect structural gaps in the literature. The analysis showed that research on heritage DTs remains fragmented across partly disconnected domains, particularly between digital representation, monitoring technologies, environmental modelling, and decision-making or governance-related issues. This fragmentation provided the empirical justification for moving beyond technology-centred maturity models toward a broader evaluative structure.

3.2 Framework proposal development

Existing DT frameworks in heritage contexts primarily address technical and structural maturity; our proposed framework extends this to encompass the social, governance, and environmental dimensions essential for sustainable heritage preservation. Drawing on Lovell et al. (2023), this research incorporates the "Usefulness" concept to ensure HBIM development aligns with stakeholder requirements. This governance and social perspective prevent overdesign by prioritizing models that fulfill specific needs without unnecessary complexity or excessive detail. The second phase translated the bibliometric findings into a multidimensional maturity framework. The derivation logic was functional rather than merely classificatory: a DT for coastal and underwater heritage should be considered mature only if it can operate as a preservation-oriented system rather than as an isolated digital model. This assumption is consistent with recent work showing that heritage DTs depend not only on 3D modelling quality, but also on interoperable workflows, semantic consistency, provenance, environmental responsiveness, and decision-support capacity (Barzaghi et al., 2024; Niccolucci et al., 2023). On this basis, five interdependent dimensions were defined. The Technical dimension addresses data acquisition, sensing, updating workflows, and analytical or predictive capabilities, and is conceptually anchored to Technology Readiness Levels (TRLs) as a reference for progression from experimental to operational systems (Mankins, 2009). The Structural dimension evaluates geometric reliability, semantic enrichment, provenance, and interoperability of the digital representation, drawing on BIM/HBIM maturity and Level of Development (LOD) principles (Underwood & Isikdag, 2010). The Environmental dimension considers the integration of coastal or underwater monitoring data, hazard variables, and predictive environmental modelling, in line with the management-oriented logic of ISO 14001 (ISO, 2015). The Governmental dimension assesses governance arrangements, stakeholder coordination, accountability, and organizational readiness, using ISO 38500 (ISO, 2024) as a conceptual anchor for governance principles. Finally, the Social dimension captures documentation quality, training, knowledge transfer, and the long-term capacity of institutions to maintain and reuse the system, consistently with the broader responsibility and stakeholder orientation reflected in ISO 2600 regarding Social Responsibility (ISO, 2010). These standards were not adopted as complete maturity models in their original form. Rather, they were used as interpretive anchors to define dimension-specific criteria while preserving a heritage-oriented perspective. This choice was intended to avoid the direct transfer of industrial or corporate maturity logic to cultural heritage, while still ensuring conceptual rigor and inter-domain comparability.

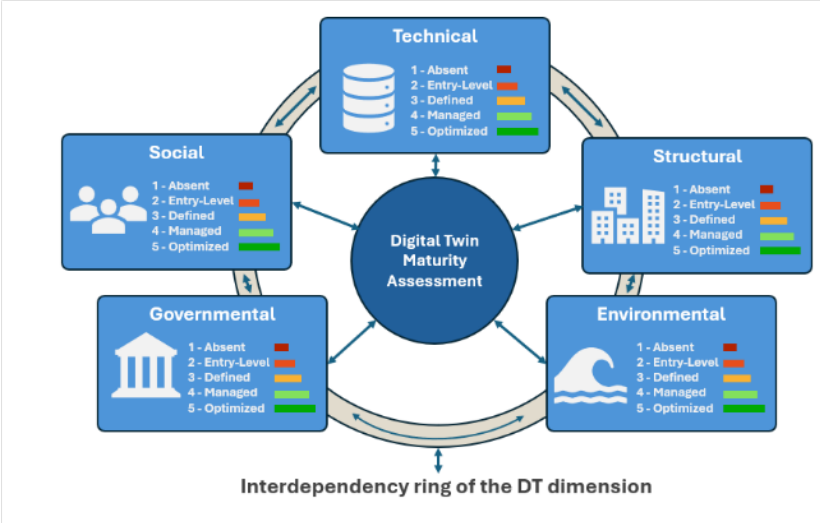


Figure 1 – DT Maturity framework assessment proposal

The proposed framework (Figure 1) operationalizes five interdependent dimensions for assessing Digital Twin (DT) maturity within the specialized context of coastal heritage preservation. This multidimensional approach moves beyond purely technical metrics to encompass the structural, environmental, governmental, and social facets of heritage management.

3.3 Maturity levels and scoring logic

Each dimension is assessed on a five-level ordinal scale (0–4): 0 = Absent, 1 = Entry Level, 2 = Defined but static, 3 = Managed, and 4 = Optimized. This scale was designed to balance simplicity and interpretability while still distinguishing between experimental, operational, and advanced DT implementations. An overall maturity score is calculated through a weighted aggregation of the five dimensions:

$$DT_{Score} = \frac{\sum_{i=1}^{5} w_i * D_i}{\sum_{i=1}^{5} w_i}$$

Where D_i represents the score of each dimension and w_i its relative weight. In the present study, equal weighting is assumed as the default condition to avoid arbitrary prioritization at this stage. However, the framework allows future calibration through expert-based methods such as the Analytical Hierarchy Process (AHP) or related multi-criteria decision-making approaches (Saaty, 1987). The aggregated score is intended as a diagnostic and comparative indicator, not as a substitute for expert judgement: the framework assumes only partial compensability across dimensions, it means that high technical or structural performance should not be interpreted as sufficient maturity when environmental integration, governance capacity, or knowledge continuity remain weak. The main output of the methodology is a multidimensional maturity assessment framework that can support the evaluation, comparison, and staged development of DT initiatives for coastal and underwater cultural heritage. Rather than ranking projects in purely technological terms, the framework is intended to identify implementation gaps and guide the transition from isolated digital products to interoperable, governable, and sustainable preservation systems

4 Findings

4.1 Bibliometric findings

Bibliometric analysis of annual publication output (2010–2025) demonstrates a significant growth (+25.63% with a strong peak in the latest years) in scientific research productivity within this domain (Figure 2).

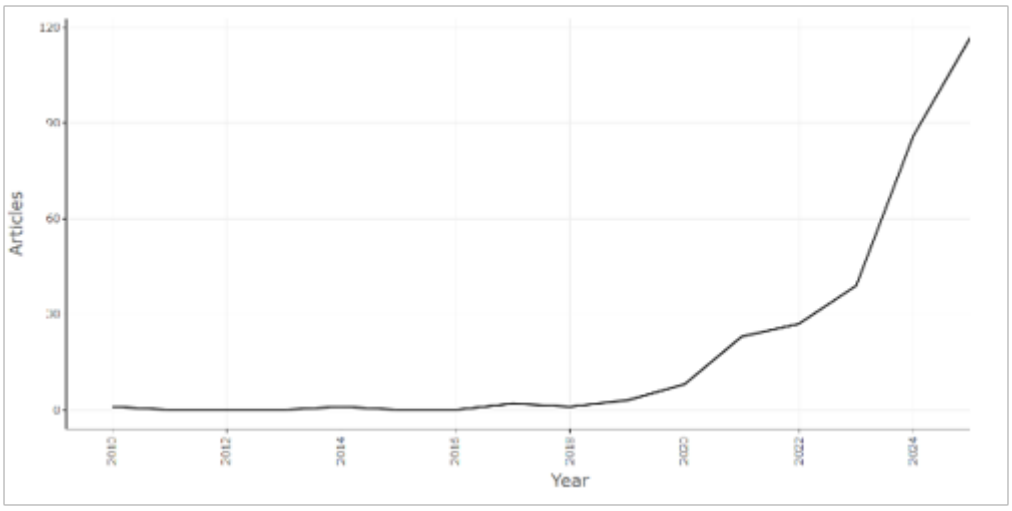


Figure 2 Annual Scientific production (Information produced by the Authors through Bibliometrix Software)

The analysis identifies three distinct temporal phases: a dormancy phase (2010–2017) characterized by minimal publication activity (average 2.1 articles/year), a growth acceleration phase (2018–2023) with gradually increasing publication rates (average 8.3 articles/year), and an emergence phase (2024–2025) exhibiting accelerated publication dynamics (peak of 120 articles in 2024). The cumulative publication growth from 2017 to 2024 represents a significant increase, indicating rapid field maturation and mainstream research adoption.

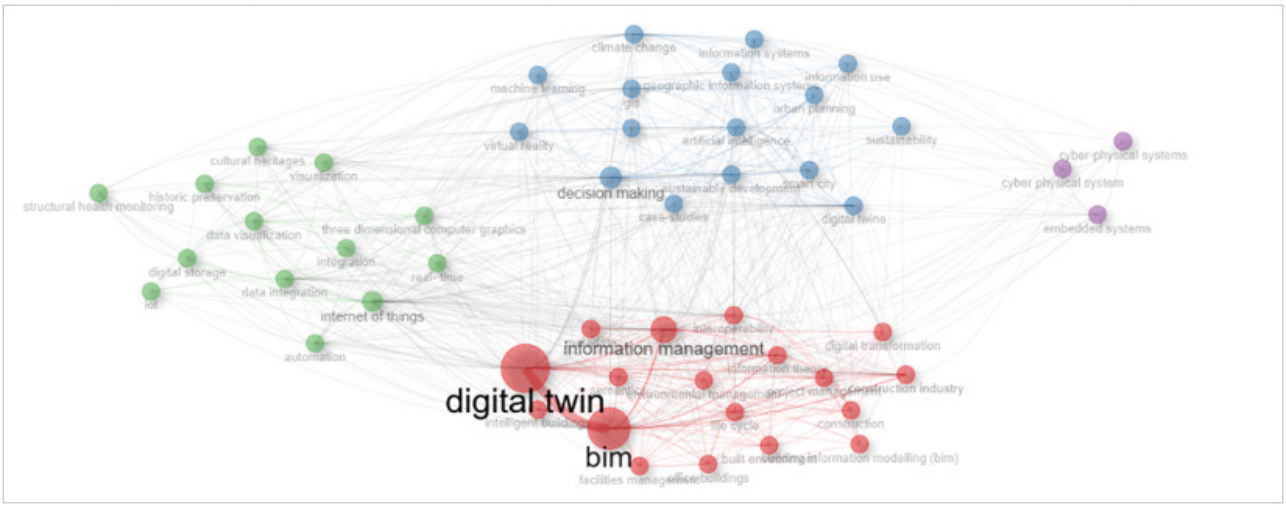


Figure 3: Keyword co-occurrence network

Keyword co-occurrence network analysis (Figure 3) demonstrates Digital Twin and BIM as dominant research nodes (red cluster, central position). Four distinct research topics emerge: information management (red), sustainability and GIS (blue), IoT and sensors (green), and Cyber-Physical Integration and Control Systems (purple). Weak inter-cluster connectivity indicates significant research fragmentation. This pattern substantiates the necessity for comprehensive maturity assessment frameworks integrating these isolated research domains. The analysis also highlights a limited overlap between core research domains. Only a small proportion of the reviewed literature explicitly combines HBIM, GIS, and Digital Twin approaches, while governance and knowledge-transfer themes remain significantly underrepresented. This finding supports the advanced argument in the previous sections: although the DT concept is gaining visibility in heritage research, its operational maturity is still constrained by fragmentation across technical systems, data environments, and institutional practices. Taken together, these bibliometric findings suggest the need for a multidimensional maturity framework. They show that the central challenge is no longer the visibility of the DT paradigm itself, but the absence of integrated evaluation criteria capable of linking representation, monitoring, environmental responsiveness, governance, and long-term usability within a single assessment structure.

4.2 Multidimensional maturity findings

Technical dimension findings

The literature shows that the Technical dimension is among the most advanced. Many projects already integrate multiple sources of digital documentation and environmental data, and several studies report the use of sensor-based acquisition, remote sensing, or interoperable data streams. In maturity terms, most current implementations can be positioned between Level 2 and Level 3, meaning that technical procedures are increasingly standardized and, in some cases, operational. However, fully automated, continuously updated, and real-time DT environments remain limited to a relatively small number of advanced research initiatives.

Structural dimension findings

The Structural dimension also shows comparatively strong development. Three-dimensional surveying, photogrammetry, laser scanning, and HBIM-based representations are now well established across a broad part of the heritage literature. As a result, many initiatives achieve Level 2–3 maturity in terms of geometric reliability and basic semantic structuring. Nevertheless, the transition from accurate digital representation to richly interconnected semantic environments remains incomplete. Provenance, ontology-based enrichment, and full interoperability with external information systems are still not systematically implemented, which limits the capacity of DTs to function as long-term knowledge infrastructures.

Environmental dimension findings

The Environmental dimension remains one of the least mature, despite being fundamental for coastal and underwater heritage protection. The reviewed literature shows that robust environmental maturity requires the integration of hydrodynamic, wave, and morphodynamic simulations able to capture water levels, currents, temperature, salinity, sediment transport, shoreline evolution, and site-specific marine forcing. Models such as ADCIRC, NEMO, ROMS, SCHISM, and SHYFEM for hydrodynamics (Fox-Kemper et al., 2019) demonstrate that high-resolution environmental forecasting is technically feasible when supported by downscaled Copernicus and EMODNET data. However, these capabilities are still only weakly integrated into most heritage DT applications, where environmental information often remains static, partial, or disconnected from structural documentation and conservation planning.

Governmental dimension findings

The Governmental dimension remains underdeveloped, even though the reviewed material shows that environmental modelling becomes operationally valuable only when translated into decision-support processes. Numerical simulations and Digital Twin environments can support adaptation planning, maintenance prioritization, and risk-informed intervention, but this requires the active involvement of public authorities, infrastructure operators, and other institutional stakeholders. In this respect, governance maturity depends not only on stakeholder presence but on the existence of procedures capable of transforming environmental evidence into coordinated action, best-practice guidelines, and policy-relevant responses. Current DT initiatives rarely demonstrate this degree of institutional embedding, which suggests that governance still lags behind technical capability.

Social dimension findings

The Social dimension shows similarly low maturity. The reviewed material suggests that Digital Twins become socially meaningful only when complex technical outputs, such as flood forecasts, erosion trends, or environmental risk simulations, are made usable for non-specialist actors and heritage practitioners. This implies not only communication and dissemination, but also documentation, training, and transferable operational knowledge. The production of best practices, replicable workflows, and accessible decision-support tools therefore becomes central to long-term sustainability. Yet these elements remain limited in current heritage DT initiatives, which still depend heavily on expert-led research settings and only rarely provide structured mechanisms for broader knowledge transfer or durable institutional learning.

5 Discussion

The findings reveal a structural paradox in the current development of heritage Digital Twins. On the one hand, the field has made clear progress in data acquisition, 3D representation, and technical modelling. On the other hand, these advances have not yet translated into equally mature systems of environmental integration, institutional governance, and knowledge continuity. This confirms that DT maturity in coastal and underwater heritage cannot be equated with modelling sophistication alone. A technically advanced system may remain immature if it cannot support risk-informed decision-making, coordinated management, and long-term operational use. From this perspective, the framework proposed in this paper helps reposition Digital Twins from a technology-centered paradigm to a preservation-centered one. Its main value lies in making explicit the dimensions that are usually treated as external or secondary to DT development. In particular, the

environmental dimension is decisive for coastal and underwater heritage because exposure to erosion, flooding, salinization, and marine dynamics directly affects the relevance of any monitoring architecture. Likewise, the governmental and social dimensions make visible the institutional and human conditions under which a DT can remain functional after the completion of a research project. A second implication concerns the interdependence of the five dimensions. The results suggest that maturity should not be understood as a linear accumulation of features, but as a relational condition in which weaknesses in one dimension constrain the system's overall usefulness. For example, strong structural modelling cannot compensate for the absence of environmental forecasting, and high technical readiness cannot substitute for weak governance or missing knowledge-transfer mechanisms. For this reason, the proposed framework encourages a parallel and coordinated development of all five dimensions from the early phases of project design. More broadly, the framework offers a practical contribution for heritage managers, archaeologists, engineers, and policymakers. It provides a common reference for identifying implementation gaps, prioritizing investment, and comparing initiatives beyond purely technological criteria. In doing so, it supports a shift from isolated digital outputs toward interoperable and sustainable systems capable of informing conservation strategies under conditions of accelerating climate risk.

6 Conclusion

This study proposed a multidimensional maturity assessment framework for Digital Twins in coastal and underwater cultural heritage. Grounded in a bibliometric analysis of 346 publications and in the synthesis of current literature, the framework was developed to address a central gap in the field: the lack of evaluative models capable of linking technical maturity with structural quality, environmental responsiveness, governance capacity, and knowledge sustainability. The results show that current DT initiatives are progressing unevenly. Technical and structural components are becoming increasingly mature, while environmental, governmental, and social dimensions remain significantly underdeveloped. This imbalance limits the practical effectiveness of DTs as tools for long-term heritage preservation, especially in coastal and underwater contexts where climate-related pressures and institutional fragmentation are particularly acute. The main contribution of the paper is therefore not only the definition of five assessment dimensions, but the reframing of DT maturity as a multidimensional preservation condition rather than as a purely technological stage. In this sense, the framework may support more consistent project design, more transparent evaluation, and more sustainable implementation pathways for initiatives such as ARCHAEOSAFE. Future research should focus on empirical validation through case studies, indicator refinement, and expert-based calibration of dimension weights.

7 Acknowledgements

This research was partially funded by the ARCHAEOSAFE project, “Archaeological Safety and Risk Management for Coastal and Underwater Heritage” (Project ID SA-0300545) funded under the Interreg VI-A IPA CBC South Adriatic Italy-Albania-Montenegro Programme.

8 Ethics Declaration

The authors declare that formal ethical clearance is not required for the article, as the research involved no human subjects and no invasive fieldwork.

9 AI Declaration

AI tools were used to support language editing and improve clarity and data preparation. The authors have reviewed and verified all AI-assisted outputs. No AI was used in the research design, analysis, or interpretation of results.

References

Aria, M., & Cuccurullo, C. (2017). bibliometrix: An R-tool for comprehensive science mapping analysis. Journal of Informetrics, 11(4), 959–975. https://doi.org/10.1016/j.joi.2017.08.007

Barzaghi, S., Bordignon, A., Gualandi, B., Heibi, I., Massari, A., Moretti, A., Peroni, S., & Renda, G. (2024). A Proposal for a FAIR Management of 3D Data in Cultural Heritage: The Aldrovandi Digital Twin Case. https://doi.org/10.48550/ARXIV.2407.02018

Chrissis, M. B., Konrad, M., & Shrum, S. (2003). CMMI: Guidelines for process integration and product improvement. Addison-Wesley.

European Commission. Directorate General for Research and Innovation. (2017). Technology readiness level: Guidance principles for renewable energy technologies : final report. Publications Office. https://doi.org/10.2777/577767

Fox-Kemper, B., Adcroft, A., Böning, C. W., Chassignet, E. P., Curchitser, E., Danabasoglu, G., Eden, C., England, M. H., Gerdes, R., Greatbatch, R. J., Griffies, S. M., Hallberg, R. W., Hanert, E., Heimbach, P., Hewitt, H. T., Hill, C. N., Komuro, Y., Legg, S., Le Sommer, J., … Yeager, S. G. (2019). Challenges and Prospects in Ocean Circulation Models. Frontiers in Marine Science, 6, 65. https://doi.org/10.3389/fmars.2019.00065

Fuller, A., Fan, Z., Day, C., & Barlow, C. (2020). Digital Twin: Enabling Technologies, Challenges and Open Research. IEEE Access, 8, 108952–108971. https://doi.org/10.1109/ACCESS.2020.2998358

Gabellone, F. (2022). Digital Twin: A new perspective for cultural heritage management and fruition. ACTA IMEKO, 11(1), 7. https://doi.org/10.21014/acta_imeko.v11i1.1085

Grazianova, M., & Mesaros, P. (2024). Cultural heritage management from traditional methods to digital systems: A review from bim to digital twin. E3S Web of Conferences, 550, 01015. https://doi.org/10.1051/e3sconf/202455001015

Gregory, D., Dawson, T., Elkin, D., Tilburg, H. V., Underwood, C., Richards, V., Viduka, A., Westley, K., Wright, J., & Hollesen, J. (2022). Of time and tide: The complex impacts of climate change on coastal and underwater cultural heritage. Antiquity, 96(390), 1396–1411. https://doi.org/10.15184/aqy.2022.115

Hou, H. (Cynthia), Lai, J. H. K., Wu, H., & Wang, T. (2024). Digital twin application in heritage facilities management: Systematic literature review and future development directions. Engineering, Construction and Architectural Management, 31(8), 3193–3221. https://doi.org/10.1108/ECAM-06-2022-0596

ISO. (2010). ISO 26000 Social responsibility. https://store.uni.com/uni-en-iso-14001-2015

ISO. (2015). UNI EN ISO 14001:2015—Sistemi di gestione ambientale—Requisiti e guida per l'uso. https://store.uni.com/uni-en-iso-14001-2015

ISO. (2024). ISO/IEC 38500:2024 Information technology—Governance of IT for the organization. https://www.iso.org/standard/81684.html

Klar, R., Arvidsson, N., & Angelakis, V. (2024). Digital Twins' Maturity: The Need for Interoperability. IEEE Systems Journal, 18(1), 713–724. https://doi.org/10.1109/JSYST.2023.3340422

Lovell, L., Davies, R., & Hunt, D. (2023). The Application of Historic Building Information Modelling (HBIM) to Cultural Heritage: A Review. Heritage, 6, 6691–6717. https://doi.org/10.3390/heritage6100350

Mankins, J. C. (2009). Technology readiness assessments: A retrospective. Acta Astronautica, 65(9–10), 1216–1223. https://doi.org/10.1016/j.actaastro.2009.03.058

Mongeon, P., & Paul-Hus, A. (2016). The journal coverage of Web of Science and Scopus: A comparative analysis. Scientometrics, 106(1), 213–228. https://doi.org/10.1007/s11192-015-1765-5

Niccolucci, F., Markhoff, B., Theodoridou, M., Felicetti, A., & Hermon, S. (2023). The Heritage Digital Twin: A bicycle made for two. The integration of digital methodologies into cultural heritage research. Open Research Europe, 3, 64. https://doi.org/10.12688/openreseurope.15496.1

Orr, S. A., Richards, J., & Fatorić, S. (2021). Climate Change and Cultural Heritage: A Systematic Literature Review (2016–2020). The Historic Environment: Policy & Practice, 12(3–4), 434–477. https://doi.org/10.1080/17567505.2021.1957264

Pavlidis, G., Sevetlidis, V., & Arampatzakis, V. (2025). Defining a New Digital Twin Ontology for Cultural Heritage Preservation—The Case of ARGUS. Digital Heritage. https://doi.org/10.2312/DH.20253260

Reimann, L., Vafeidis, A. T., Brown, S., Hinkel, J., & Tol, R. S. J. (2018). Mediterranean UNESCO World Heritage at risk from coastal flooding and erosion due to sea-level rise. Nature Communications, 9(1), 4161. https://doi.org/10.1038/s41467-018-06645-9

Saaty, R. W. (1987). The analytic hierarchy process—What it is and how it is used. Mathematical Modelling, 9(3–5), 161–176. https://doi.org/10.1016/0270-0255(87)90473-8

Underwood, J., & Isikdag, U. (Eds.). (2010). Handbook of Research on Building Information Modeling and Construction Informatics: Concepts and Technologies. IGI Global. https://doi.org/10.4018/978-1-60566-928-1

UNESCO. (2001). Convention on the Protection of the Underwater Cultural Heritage. https://www.unesco.org/en/legal-affairs/convention-protection-underwater-cultural-heritage

Artificial Intelligence and the Transformation of Knowledge Management Systems: A Consensus Community-Detection Bibliometric Analysis

Carlo Drago[1], Angelo Leogrande[2], Valentina Giglio Moro[1], Linda Stella Guajardo Rios[1]

[1] University Niccolò Cusano, Rome, Italy

[2] LUM University, Giuseppe Degennaro, Casamassima (BA), Italy

Abstract

The integration of Artificial Intelligence (AI) and Generative Artificial Intelligence (GenAI) is transforming Knowledge Management. Digital technologies are evolving from passive information repositories into active knowledge co-constructors capable of supporting knowledge generation, recombination, dissemination, interpretation, and application. This study examines the emerging relationship among Knowledge Management, AI, and GenAI, focusing on their implications for firms operating in increasingly data-intensive, uncertain, and knowledge-driven environments. A key difference exists between AI-enabled Knowledge Management and current systems. As technology and society evolve, businesses learn, make decisions, generate new ideas, compete, and govern in new ways. This research employs a bibliometric methodology centered on Scopus data, keyword co-occurrence networks, and ensemble MCA-based community detection techniques. A consensus community detection method combines multiple algorithms and merges their outputs using dimensionality reduction and hierarchical clustering to enhance methodological rigor. This approach allows the analysis to move beyond reliance on a single algorithm, identifying more stable thematic structures within the literature. The objective is to uncover the principal thematic cores, novel research trajectories, and uncharted gaps in AI-enabled knowledge creation, dissemination, storage, application, decision support, organizational performance, ethical governance, trust, explainability, and human–AI collaboration. This study contributes to the literature on AI-enabled Knowledge Management by providing a systematic conceptual mapping of a rapidly evolving field. It also offers practical insights for managers and organizations seeking to understand how AI and GenAI can reshape knowledge processes while generating new challenges related to data quality, algorithmic bias, privacy, accountability, organizational readiness, employee skills, and responsible implementation. In this context, AI and GenAI are interpreted as transformative infrastructures through which firms may redesign knowledge ecosystems, augment human expertise, and develop more adaptive, intelligent, and strategically oriented knowledge management capabilities. By examining AI's potential for boosting creativity, automating tasks, and improving decision-making, we also take into consideration ethical issues like fairness and inclusion. By balancing innovation with responsible practices, businesses can achieve a healthy balance. In a world that is becoming increasingly AI-driven, companies are making sure AI allows them to drive sustainable growth, give employees more power, and fit with long-term strategic goals.

Keywords: Artificial Intelligence, Knowledge Management, Generative AI, Bibliometric Analysis, Ensemble Community Detection

Paper type: Academic Research Paper

1 Introduction

In the rapidly evolving digital landscape, Knowledge Management (KM) is undergoing a fundamental paradigm shift. Previously, the field was characterized by human-centric curation and static databases, but today it is being transformed by the integration of Artificial Intelligence (AI) and Generative AI (GenAI). This transition marks a move away from viewing technology as a passive repository toward recognizing it as an active knowledge "co-constructor" capable of coexisting with human experts (see Sun and Xu 2025). This significant change requires close attention. AI is now structurally integrated into the very definition of strategic knowledge management, as illustrated by current literature (Yuan et al. 2024).

A recent industry survey reported that many organizations now use AI in at least one business function (McKinsey & Company, 2025) as GenAI becomes embedded in workflows. Organizations are improving their ability to create, store, share, and apply knowledge by leveraging Large Language Models (LLMs) and advanced natural language processing (NLP) (Nakash and Bolisani 2025). In KM, AI is perceived as highly beneficial across all processes, especially in knowledge acquisition and documentation, where it significantly reduces cognitive load and streamlines data processing (Nakash and Bolisani 2025). However, technological integration poses several challenges. As part of the implementation process, organizations must restructure their infrastructure, talent, and workflows (Singla et al. 2025). At the same time, these changes are compounded by ethical dilemmas

related to data privacy, algorithmic bias, and ownership, as well as employee resistance due to fears of job displacement and role ambiguity (see Leoni et al. 2024; Rezaei 2025; Thakuri et al. 2024). There are also significant financial and technical barriers, including integration with legacy systems, skills gaps, and resource constraints (Rezaei 2025).

Integrating Artificial Intelligence into Knowledge Management is challenging and requires a comprehensive approach that considers multiple strategies. To maximize AI's potential without compromising organizational integrity, companies should focus on three key areas: establishing robust governance frameworks, implementing ongoing upskilling programs, and executing strategic change management. Robust governance frameworks are essential for integrating AI into KM, as they ensure responsible and ethical use (see Leoni et al. 2024). Overall, these processes should not be viewed solely as a technical responsibility of IT departments.

The research question in this work explores, using bibliometric analysis, the relationship between knowledge management, artificial intelligence, and generative AI, with the aim of understanding how they can impact firms.

The work is organized as follows: in the second section, we consider the data used and the methodology adopted for the bibliometric analysis. In the third section, we discuss the results obtained. In the fourth section, we present the conclusions of the work.

2 Data and Methodology

To investigate these phenomena, we will use a bibliometric analysis approach based on co-occurrence networks and ensemble community detection (see Drago 2018, Drago & Errichiello 2026, Morelli et al. 2026 and Fortunato 2010). Community detection was applied to the relevant group of keywords identified in the literature as the core of the field. We used appropriate queries related to the literature to explore the field, and using the results from these queries in Scopus, we constructed the bibliometric dataset. In this respect, the query used was:

TITLE-ABS-KEY(
("knowledge management system" OR KMS OR "knowledge management" OR "organizational knowledge")*
AND
("artificial intelligence" OR AI OR "machine learning" OR "deep learning" OR "generative AI" OR "large language model" OR LLM* OR "natural language processing" OR NLP)*
AND
(integrat OR implement* OR adopt* OR deploy* OR "socio-technical" OR "human-in-the-loop" OR "hybrid intelligence" OR augment* OR barrier* OR challenge* OR limitation* OR risk* OR govern* OR ethic* OR bias OR fairness OR transparen* OR explainab* OR interpretab* OR accountab* OR "knowledge quality" OR valid* OR reliab* OR "knowledge verification" OR "knowledge validation" OR trust*)*
)

The query was performed on January 18, 2026. From the bibliometric dataset, we conducted the bibliometric analysis using R (see Aria & Cuccurullo 2017). The analysis focused on the co-occurrence keyword network. To improve the robustness of the results and avoid potential bias from any single algorithm (see Leskovec et al. 2010), we used a consensus approach. Consensus community detection helps synthesize the results for the different membership keywords by community (Drago & Balzanella 2015) for each method. Following the consensus clustering algorithm proposed by Drago (2018), we applied a multidimensional reduction technique, such as multiple correspondence analysis (MCA; see Greenacre & Blasius 2006), to reduce the dimensionality of this matrix. Validation was performed using different criteria to define the optimal number of clusters in the hierarchical clustering process of the keywords (Brock et al. 2008 on clustering validation). Finally, we obtained the relevant groups of keywords that represent the most significant cores (and keyword communities in the sense of Fortunato 2010) of the literature considered (see Gatto et al. 2023, Drago et al. 2023).

3 Results and Discussion

From the bibliometric analysis, we obtained the dendrogram presented in Fig. 1. The results of the final dendrogram can be analyzed cluster by cluster, or by community, considering that this is a community detection procedure, as follows:

Cluster 1, Artificial Intelligence and Knowledge Management keywords can be seen as the foundational conceptual core of recent literature bon AI-KM integration. It is not simply a collection of contributions on adopting new technologies, but rather identifies a genuine transformation of the knowledge management paradigm: from predominantly human, document-based, and repository-based systems to dynamic, predictive,

adaptive, and partly generative systems, in which artificial intelligence intervenes in the processes of identifying, acquiring, creating, archiving, sharing, and applying knowledge.

The central logic of the cluster is that AI is increasingly viewed as a cognitive and organizational infrastructure capable of extending the traditional capabilities of knowledge management. From this perspective, AI is not merely a technical tool but a socio-technical device that automates certain knowledge-intensive activities, enhances the capacity to analyze large volumes of data, facilitates the extraction of latent patterns, and supports faster and more informed decision-making. Leoni et al. (2024) emphasize that knowledge management processes involve the identification, acquisition, creation, storage, sharing, and application of knowledge, and that AI enables the knowledge spectrum to be broadened, making it accessible to all people, in every situation and at any time. In bibliometric terms, this cluster can therefore be interpreted as the foundational domain in which the AI–knowledge management relationship is conceptualized as a convergence between algorithmic capability, organizational knowledge management, and decision-making.

The keywords for the **second cluster** are: large language models, generative AI, ChatGPT, LLM, and LLMs. Cluster 2 could be defined as "Generative AI and Large Language Models for Knowledge Work." This cluster identifies the most recent and specific component of the literature on the integration of artificial intelligence and knowledge management. While the first cluster addressed the general relationship between AI and KM, this second cluster focuses on the new generation of generative language technologies, particularly Large Language Models and tools such as ChatGPT, which are considered capable of transforming everyday knowledge work. Conceptually, the cluster marks a shift from predominantly analytical, predictive, and classificatory AI to generative, conversational, and interactive AI. LLMs do not merely extract information from large databases; they produce texts, summaries, explanations, drafts, decision-making scenarios, and knowledge artifacts. Bratianu and Ioan (2025) note that Generative AI, as a subfield of AI, uses natural language processing, neural networks, and machine learning to create human-like content, and that Large Language Models process enormous datasets comprising natural languages, code, mathematical expressions, and images. ChatGPT is cited as one of the best-known applications of this technological family.

The significance of this cluster is closely linked to the transformation of organizational knowledge into knowledge that can be generated, summarized, and reused through linguistic interaction. LLMs make knowledge work more accessible even to non-specialist users, as they allow repositories, documents, and information bases to be queried using natural language. Rezaei (2025) emphasizes that tools such as ChatGPT support organizations in generating personalized content, summarizing complex documents, retrieving context-specific knowledge, and personalizing interactions with users.

The keywords of the **third cluster** are: network exposure; NaaS (Network as a Service); ml driven-analytics; intent-based networking; entreprise industries; end-to-end automation; disaggregation; data lake; cognitive networks; cloud platform; cloud native; cloud; closed loop operations; autonomous networks; automation transparency; analytics; edge computing; automation; data management; operational efficiency. This third cluster can be interpreted as representing intelligent, cloud-native, and automated infrastructure for organizational knowledge management. Unlike the group focused on the linguistic production of content, such as large language models and generative AI, this cluster centers on the technical and organizational architecture that enables advanced enterprise systems to collect, process, integrate, circulate, and use knowledge. The keywords indicate an environment where knowledge is no longer treated merely as a document, repository, or encoded information, but as a continuous operational flow embedded in networks, cloud platforms, autonomous systems, data lakes, edge computing, and end-to-end automation mechanisms.

The joint presence of cloud, cloud platform, cloud-native, data lake, and data management indicates that this cluster addresses the infrastructural dimension of knowledge management systems. From this perspective, organizational knowledge is increasingly embedded in distributed, scalable, and interoperable architectures, where data from multiple sources is stored, integrated, and made available for analytical and decision-making functions. This is consistent with recent literature, which states that AI enables organizations to automate the management of large volumes of information, improve knowledge search and retrieval, classify documents, and integrate heterogeneous sources into more cohesive and usable knowledge bases. In particular, Rezaei (2025) notes that AI enhances knowledge storage by automating data organization, classification, search, and integration of diverse sources, affecting the accessibility, performance, efficiency, and security of knowledge systems.

The keywords of the **fourth cluster** are: ontology; AI; knowledge management; knowledge sharing; artificial intelligence; digital transformation; knowledge transfer; innovation; risk management; blockchain; trust; tacit knowledge; digitalization; sustainability; Industry 4.0. This cluster outlines a research trajectory focused on the semantic, socio-technical, and strategic dimensions of AI in knowledge management systems. It highlights that integrating artificial intelligence into KMSs requires not only algorithmic capabilities but also ontological

structures to represent knowledge, organizational processes to facilitate sharing and transfer, governance mechanisms to manage risks and reliability, and a persistent centrality of the human element, especially in processes involving tacit knowledge, interpretation, and managerial judgment.

In summary, the fourth cluster can be interpreted as the cluster of AI-enabled knowledge governance. It demonstrates that the digital transformation of knowledge management involves not only automating archives or generating content, but also designing intelligent, semantic, and responsible systems capable of linking human and artificial knowledge, industrial innovation, knowledge transfer, sustainability, and risk management.

The fifth cluster is centered on knowledge graph and transfer learning. This cluster identifies an emerging trend in which AI is used not only to generate or automate knowledge, but also to structure it semantically and make it transferable across different organizational domains. The co-occurrence of knowledge graphs and transfer learning suggests a shift toward knowledge management systems that represent entities, concepts, and relationships in a networked form and reuse knowledge models learned in one context to address problems in new contexts. This cluster thus serves as a bridge between knowledge representation and organizational learning, showing how AI-powered KMSs can evolve from static repositories into dynamic cognitive infrastructures that integrate, transfer, and recombine knowledge. However, this trajectory presents significant challenges, particularly regarding scalability, semantic interoperability, data quality, adaptation to heterogeneous datasets, and consistency between artificial and human organizational knowledge.

The keyword of the **sixth cluster** is based on a single keyword large language models where the **seventh cluster** is based on: machine learning; deep learning; natural language processing; big data; reinforcement learning. This data structure, with large language models as the sixth cluster, suggests that this cluster represents a recent specialization closely linked to the algorithmic core of AI, but not entirely overlapping with it. In bibliometric terms, the dendrogram appears to represent a conceptual progression: at the base is the general paradigm of machine learning; from it emerge deep and linguistic methods; finally, large language models appear as a more recent and distinct frontier, derived from the convergence of deep learning, NLP, and large databases. This seventh cluster represents the core methodological and algorithmic aspect of this literature. Unlike clusters focused on organizational systems, semantic governance, or large language models, the focus here is on the fundamental techniques that enable artificial intelligence in knowledge management: machine learning, deep learning, natural language processing, and reinforcement learning. In other words, this cluster describes the "computational" level of AI-KM – the level at which data is transformed into patterns, patterns into operational knowledge, and operational knowledge into decision support.

The seventh cluster identifies the core literature on the integration of AI and knowledge management. It highlights the role of machine learning and deep learning techniques in extracting patterns from large volumes of data, managing unstructured knowledge, and processing natural language as the primary means of accessing organizational knowledge. At the same time, it is possible to deduce that these techniques, by their nature, are primarily interpreted as supporting decision-making processes, while "reinforcement learning" suggests a trajectory toward adaptive systems capable of learning from feedback, outcomes, and subsequent interactions. Therefore, the cluster can be interpreted as the domain of the algorithmic foundations of AI-enabled knowledge decision systems, in which the value of knowledge management derives from the ability to transform heterogeneous data into actionable knowledge and more timely, accurate, and adaptive organizational decisions.

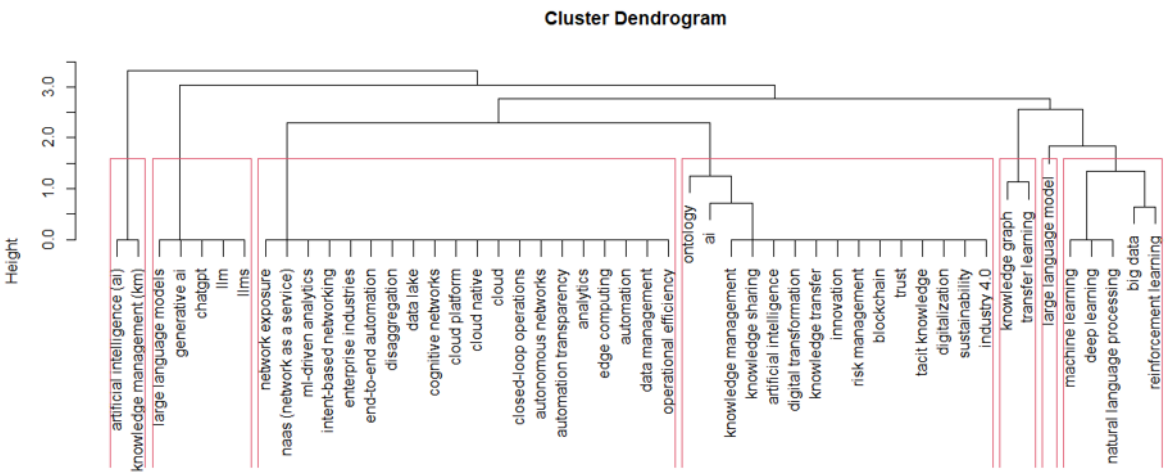


Figure 1: The final dendrogram from the hierarchical cluster analysis

4 Conclusions

This research investigated the correlation among Artificial Intelligence (AI), Generative AI (GenAI), and Knowledge Management (KM) using a bibliometric and community-detection methodology with a Scopus-based dataset. The findings indicate that the intersection of AI and KM has developed into a significant research area, highlighting the interdependence of knowledge processes, algorithmic systems, and organizational decision-making. AI-enabled KM involves more than simply adopting new tools; it represents a relevant shift in how organizations manage knowledge, moving from human-curated models to hybrid socio-technical architectures where human expertise and AI systems collaborate.

Key findings show that the AI-KM nexus is highly important, with AI serving as an enabling infrastructure for knowledge-intensive work. The rise of GenAI and Large Language Models (LLMs) demonstrates a shift toward interactive systems that generate knowledge artifacts, underscoring the need for further research on the reliability and governance of AI-generated knowledge. The research also indicates that digital infrastructures are crucial in knowledge management, suggesting that knowledge is continuously generated through operational flows and automated analytics.

Additionally, the socio-technical and governance dimensions of AI in KM are critical, requiring both human judgment and governance mechanisms alongside automation. The methodological foundations of AI-enabled KM are rooted in machine learning and natural language processing.

The study, consistent with the objectives of bibliometric research, identifies several research gaps, including the need for rigorous indicators to measure the value of AI-enabled KM, methods for validating AI-generated knowledge, and comprehensive governance frameworks for hybrid human-AI systems. It emphasizes that organizations should approach AI-enabled KM as an organizational transformation rather than a simple technological project, integrating AI into defined knowledge processes while ensuring human involvement in critical decision-making. The findings enhance understanding of the evolving landscape of AI-enabled KM and suggest directions for future research to address current gaps and improve organizational practices.

These groups of keywords can be used to identify emerging trends and research gaps in the field. They can also assist in developing targeted literature reviews, guiding researchers toward the most impactful studies. Furthermore, they may help academics and practitioners identify potential areas for collaboration and innovation. Thus, the final expected results can be used not only to understand the existing literature on this topic but also to better identify relevant research gaps that can be addressed in future studies.

5 Acknowledgements

We would like to thank the two anonymous reviewers for their valuable comments and constructive suggestions, which have helped us improve the quality and clarity of the manuscript.

6 Ethics declaration

All data used in the analysis refer to publicly indexed scientific publications and were processed in aggregate form for bibliometric and network-analytical purposes. Therefore, no institutional ethics approval was required.

7 AI declaration

AI was used only to improve the readability and language of the work. This technology was used with control, and the results were carefully reviewed. We are ultimately responsible and accountable for the contents of the work.

References

Aria, M., & Cuccurullo, C. (2017). bibliometrix: An R-tool for comprehensive science mapping analysis. Journal of informetrics, 11(4), 959-975.

Bratianu, C. and Ioan, A., 2025. Challenges of artificial intelligence for knowledge management systems: A bibliometric analysis perspective. Oradea Journal of Business and Economics, 10(1), pp.108–120.

Brock, G., Pihur, V., Datta, S., & Datta, S. (2008). clValid: An R package for cluster validation. Journal of statistical Software, 25, 1-22.

Drago, C. (2018). MCA-based community detection. In Classification,(big) data analysis and statistical learning (pp. 59-66). Cham: Springer International Publishing.

Drago, C., & Balzanella, A. (2015). Nonmetric MDS consensus community detection. In Advances in Statistical Models for Data Analysis (pp. 97-105). Cham: Springer International Publishing.

Drago, C. and Errichiello, L. (2026) 'Remote work during pandemics: a systematic review based on keyword-ensemble community detection', *Quality & Quantity*. doi: 10.1007/s11135-026-02700-6.

Fortunato, S. (2010). Community detection in graphs. Physics reports, 486(3-5), 75-174.

Gatto, A., Drago, C., & Ruggeri, M. (2023). On the frontline—a bibliometric study on sustainability, development, coronaviruses, and COVID-19. Environmental Science and Pollution Research, 30(15), 42983-42999.

Greenacre, M., & Blasius, J. (2006). Multiple correspondence analysis and related methods. Chapman and Hall/CRC.

Leoni, L., Gueli, G., Ardolino, M., Panizzon, M. and Gupta, S. 2024. 'AI-empowered KM processes for decision-making: empirical evidence from worldwide organisations', *Journal of Knowledge Management*, 28(11), pp. 320–347.

Leskovec, J., Lang, K. J., & Mahoney, M. (2010, April). Empirical comparison of algorithms for network community detection. In Proceedings of the 19th international conference on World wide web (pp. 631-640).

McKinsey & Company. (2025). The state of AI: Global survey 2025. https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai

Morelli, G., Gurrieri, A.R., Drago, C., Gatto, A. and Magazzino, C., 2026. A biblio-ecological study on sustainable land planning research: evidence from machine learning methods. *Socio-Economic Planning Sciences*, article 102431. doi: 10.1016/j.seps.2026.102431.

Nakash, M. and Bolisani, E., 2025. The transformative impact of AI on knowledge management processes. Business Process Management Journal, 31(8), pp.124-147.

Rezaei, M. 2025. 'Artificial intelligence in knowledge management: Identifying and addressing the key implementation challenges', *Technological Forecasting and Social Change*, 217, 124183.

Singla, A., Sukharevsky, A., Yee, L., Chui, M. and Hall, B., 2025. The state of AI. How Organizations are Rewiring to Capture Value. Publisher: McKinsey.

Sun, Y. and Xu, C., 2025. Hybrid cognitive authority and algorithmic subjectivity: rethinking knowledge management in AI-driven communication. Journal of Knowledge Management, pp.1-26.

Thakuri, S., Bon, M., Cavus, N. and Sancar, N. 2024. 'Artificial Intelligence on Knowledge Management Systems for Businesses: A Systematic Literature Review', *TEM Journal*, 13(3).

Yuan, D., Kim, J., & Cong, G. (2024). Adoption of Artificial Intelligence and Its Impact on Competitive Advantage: Mediated by Knowledge Management. Journal of Information & Knowledge Management, 24(02). https://doi.org/10.1142/s0219649225500030

Governing Adaptation: Family Councils, Dynamic Capabilities, and Company Performance in Family Firms

Attila Wieszt, Noémi Szilvia Lőrincz, Amitabh Anand
Corvinus University of Budapest, Budapest, Hungary

Abstract

Family firms often rely on resources such as long-term orientation, trust, patient capital, ownership commitment and accumulated knowledge, yet these resources may generate rigidity, risk aversion and resistance to renewal. This paper examines how this paradox is shaped by family councils, understood as governance forums operating at the boundary of family, ownership and business. While family governance, dynamic capabilities and family firm performance have each received scholarly attention, their interrelationship remains underdeveloped. In particular, little is known about how family council characteristics and functioning influence the development of dynamic capabilities and condition their translation into company performance.

The paper develops a review-based conceptual model through a structured integrative literature synthesis. It connects three literature streams: family council and family governance research, dynamic capabilities research in family firms, and studies on family firm performance. Rather than presenting a systematic review, the study uses these streams to build a causal conceptual model of the family council–dynamic capabilities–company performance relationship.

The model proposes that family council characteristics and functioning influence the development of sensing, seizing and transforming capabilities. This relationship may be mediated or moderated by variables such as socioemotional wealth balancing, family social capital, intergenerational knowledge transfer, legitimacy creation, conflict processing and ambidextrous governance. Second, the model argues that family councils moderate the relationship between dynamic capabilities and company performance by shaping whether adaptive capabilities receive family support, ownership legitimacy and resource commitment. Third, the model includes recursive feedback loops: dynamic capabilities and performance outcomes may reshape the characteristics, legitimacy and functioning of the family council over time.

The paper contributes by positioning the family council not merely as a communication body, but as a governance infrastructure for firm-level adaptation. It also identifies future research directions on the family council–dynamic capabilities link, the dynamic capabilities–performance black box, contextual conditions and feedback processes.

Keywords: family council; dynamic capabilities; family governance; company performance; family firms

Paper type: Academic Research Paper

1 Introduction

Family firms often show the same duality. On the one hand, these firms can build on resources that are difficult for many non-family firms to access: long-term orientation, trust, patient capital, strong ownership commitment, accumulated industry knowledge, and stable relationship networks. On the other hand, these same resources can easily become rigidity. The protection of tradition may hinder renewal, the preservation of family control may slow down decisions, conflict avoidance may postpone necessary changes, and past successes may reinforce those routines that no longer work in a new market situation. The strategic question for family firms is therefore not simply whether they possess specific family resources, but whether they are able to transform these into firm-level capabilities that support adaptation and renewal.

This duality makes the question of how family firms are able to adapt particularly important. The theory of dynamic capabilities offers a strategic management explanation for this. Dynamic capabilities describe how the firm senses the opportunities and threats appearing in its environment, how it mobilizes resources to seize them, and how it transforms its resources, routines, and operations under changing circumstances (Teece, Pisano and Shuen, 1997; Teece, 2007). In the case of family firms, however, these capabilities cannot be interpreted merely as management routines. The family's goals, values, ownership logic, conflicts, and long-term intentions directly shape what the firm perceives, what it is willing to allocate resources to, and what kind of change it considers acceptable. Thus, dynamic capabilities in family firms appear as adaptation processes saturated with family meanings and ownership preferences.

This is where the family council comes in. The family council is one of the central institutions of family governance: a forum that operates at the boundary of the family, ownership, and the firm, and in which family members discuss and work through shared issues, ownership intentions, values, conflicts, and strategic questions (Gersick et al., 1997; Suess, 2014; Chrisman et al., 2018). The family council is therefore not simply a "communication space." It becomes strategically significant when it is able to organize family intentions, values,

and tensions in a way that also influences the firm's adaptive capacity. In this sense, the family council may operate as a translating and legitimizing mechanism between the family system and corporate strategic action.

In the literature, however, this relationship has not yet been sufficiently clarified. Research on family governance says much about how the family organizes its own involvement in the firm. The literature on dynamic capabilities says much about how the firm adapts in a changing environment. Research on family firm performance has long debated what performance exactly means in the case of firms that do not pursue exclusively financial goals. These research areas, however, often run alongside each other. We know little about how the characteristics and functioning of the family council contribute to the development of dynamic capabilities, and even less about how the family council influences the transformation of dynamic capabilities into company performance. The gap is therefore twofold: it is not clear how the family council becomes an antecedent of dynamic capabilities, nor under what conditions it strengthens or weakens the performance effect of dynamic capabilities.

This study aims to address this gap. We do not argue that the family council is the only factor that shapes the adaptation of family firms. Ownership structure, generational stage, family culture, firm size, the industry environment, and the intensity of competition are also important. The focus of the study is narrower: we examine how the characteristics and functioning of the family council may be connected to dynamic capabilities, and through them to company performance. Therefore, we interpret the effect of the family council not as an automatically positive direct effect, but through mechanisms, conditions, and feedback loops.

Accordingly, our research question is the following: How and under what conditions do the characteristics and functioning of the family council influence the development of dynamic capabilities, as well as their transformation into company performance in family firms?

The study is based on a structured integrative literature synthesis. We connect the results of three streams of literature: research on the relationship between the family council and dynamic capabilities, the family firm literature examining the relationship between dynamic capabilities and company performance, and works on the interpretation of family firm performance. The aim is not to present a comprehensive systematic literature review, but to develop a causal conceptual model that arranges the lessons of the three areas into a single explanatory framework.

According to the proposed model, the characteristics and functioning of the family council first influence the development of dynamic capabilities. This relationship may operate through mediating and moderating variables, for example through the balancing of socioemotional wealth-related considerations, family social capital, intergenerational knowledge transfer, legitimacy creation, conflict processing, or the support of ambidexterity. Second, the family council does not only shape the development of dynamic capabilities, but also whether these capabilities are able to become actual company performance. Here, the family council appears as a moderator: it influences whether the sensing, seizing, and transforming processes receive family support, ownership legitimacy, and sustained resource commitment. Third, the model also assumes feedback loops: the emerging dynamic capabilities and the results of company performance feed back over time into the characteristics, legitimacy, and functioning of the family council. In this way, the study proposes not a linear, but a recursive governance–capability–performance logic.

The main contribution of the study is therefore that it interprets the family council not merely as a family governance forum, but as one possible governance infrastructure of firm-level adaptive capacity. In this way, it provides a more precise picture of how, in family firms, family-level intentions and relational patterns may become firm-level capabilities — or, conversely, how they may hinder their development and transformation into performance. Overall, the strategic significance of the family council therefore does not lie in its existence in itself, but in whether it is able to place the energies, tensions, and long-term preferences of the family system in the service of firm-level renewal.

2 Methodology

This study is not a comprehensive systematic literature review in the traditional sense. The reason for this is not technical, but rather substantive. The relationship under examination—family council, dynamic capabilities, and corporate performance—is a research intersection on which few direct studies are based. If we were to consider only those articles that simultaneously use the concepts of family council, dynamic capabilities, and performance, we would not obtain a sufficient empirical and theoretical foundation for constructing a meaningful model. The relevant knowledge is therefore not found in a single, well-defined body of literature, but rather across multiple, partially overlapping research areas.

For this reason, the study employs a structured integrative literature synthesis. The aim of integrative literature synthesis is not merely to compile existing findings, but to reorganize them within a theoretically

meaningful explanatory framework (Webster and Watson, 2002; Snyder, 2019). In the case of this study, this is particularly warranted because the three central concepts are linked to different scholarly traditions. The family council appears primarily in the literature on family governance. Dynamic capabilities originate from strategic management and are subsequently applied in the context of family businesses. Performance, meanwhile, poses a distinct challenge in family business research, as family, ownership, and non-financial goals come into play alongside financial performance.

The synthesis is based on three strands of literature. The first comprises research examining the relationship between the family council and dynamic capabilities, as well as studies from which such a relationship can be indirectly inferred. In this area, the number of direct findings was low, so the search was expanded to include the relationship between family business characteristics and dynamic capabilities. This literature helps identify the family, ownership, and governance factors that may influence the development of sensing, seizing, and transforming capabilities. The second line of literature addresses the relationship between dynamic capabilities and firm performance in family businesses. We use this literature with caution because several identified mediating variables—such as innovation, knowledge integration, strategic flexibility, or entrepreneurial orientation—partially overlap with the concept of dynamic capabilities. Therefore, we do not view these as independent, clear-cut mediators, but rather as conceptually mixed constructs that nuance the interpretation of the DC–performance relationship. The third strand addresses the concept of family business performance. Its primary role is not to incorporate every possible dimension of family performance into the model, but rather to clarify that the focus of this study is on business performance, particularly performance related to adaptability, operational efficiency, market responsiveness, financial performance, and resilience.

The methodological logic of the study therefore consists of three steps. First, we identify which family council characteristics and operational features in the literature are directly or indirectly linked to dynamic capabilities. Second, we identify the variables that may mediate or moderate the relationship between the family council and dynamic capabilities. Third, we incorporate the relationship between dynamic capabilities and organizational performance into the model in a way that avoids conceptual circularity: we do not automatically treat as separate mediators those phenomena that may themselves be part of or direct manifestations of dynamic capabilities.

This approach, of course, does not replace subsequent empirical testing. The aim of the study is not to prove causal relationships, but to propose a conceptual model that more accurately maps out the possible relationships between family council, dynamic capabilities, and corporate performance. Accordingly, the results are presented not as definitive empirical statements, but as theoretically grounded relationships that can be examined in future research.

3 Findings - The Integrated FC–DC–Corporate Performance Model

Based on the literature review, the relationship between the family council, dynamic capabilities, and corporate performance cannot be described as a simple linear causal chain. The family council is not merely a precursor to dynamic capabilities, nor is it a simple mediator between dynamic capabilities and performance. In the model (see Figure 1), the family council plays a dual role. On the one hand, the characteristics and functioning of the family council shape the development of dynamic capabilities. On the other hand, these same characteristics and operational features also influence whether the dynamic capabilities that are already developing can translate into corporate performance. The model therefore focuses not on the mere existence of the family council, but on its quality, functioning, and context.

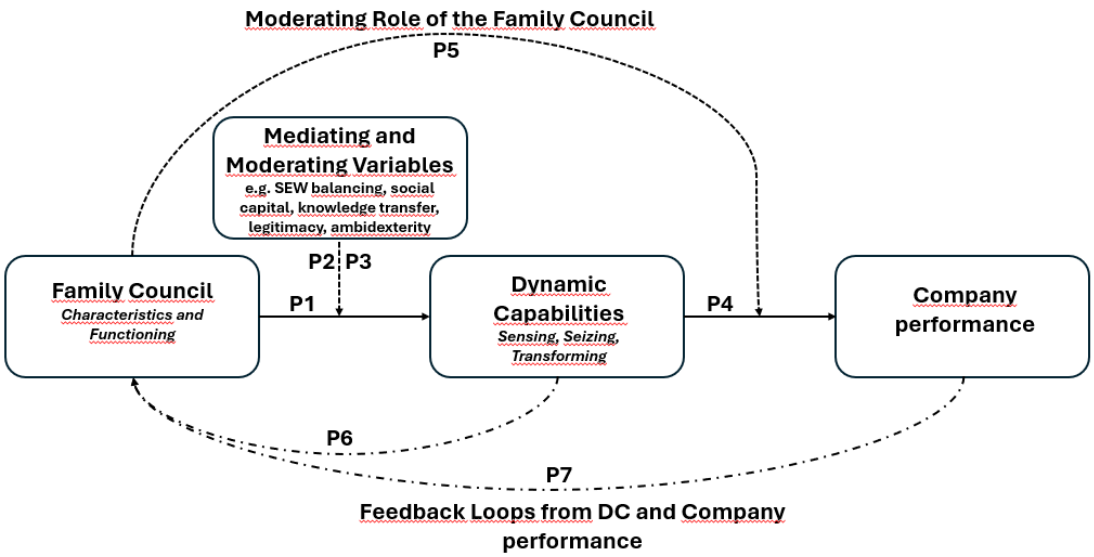


Figure 1 – The Conceptual Model (Family Council, Dynamic Capabilities and Company Performance) and our propositions

3.1 The characteristics and functioning of the family council as precursors to dynamic capabilities

The first major relationship exists between the family council and dynamic capabilities. By characteristics of the family council, we mean the relatively stable structural and formal factors that determine the council's composition, mandate, and connections. These include, for example, generational composition, representation of family branches, formal regulation, authority, decision-making or advisory roles, the regularity of meetings, and how the council relates to shareholder forums, the board of directors, or company management.

The functioning of the family council, on the other hand, refers to how the council operates in practice. Here, the quality of dialogue, trust, the ability to resolve conflicts, strategic alignment, decision-making discipline, openness, and the handling of differing opinions become important. This distinction is important because a family council can be well-designed in formal terms but function poorly. Similarly, a less formalized family council can be strong if it is actually capable of building trust, shared understanding, and a strategic direction.

The three core elements of dynamic capabilities—sensing, seizing, and transforming—can relate to the family council in different ways. Sensing, or the perception of opportunities and threats, is particularly sensitive to who participates in the family council and what information enters the family decision-making process. Generational diversity, the involvement of external advisors, the participation of younger family members, or the emergence of different corporate roles can broaden what the family perceives from its environment. If the family council is closed, homogeneous, or merely echoes the opinions of dominant family members, then sensing may narrow. In such cases, the family does not perceive environmental changes more accurately, but rather reinforces its own previous interpretations.

Seizing, or seizing opportunities, requires different types of family council conditions. Here, the legitimacy of decisions, commitment of resources, owner consensus, and the interpretation of risks from a family perspective become important. A well-functioning family council can help ensure that the family not only recognizes a strategic opportunity but is also willing to allocate resources to it. This is particularly important in family businesses, where investment decisions are not merely financial calculations but often involve identity, control, legacy, and family reputation. In this sense, the family council does not replace corporate management, but it can influence whether lasting owner support is established behind corporate management.

Transformation—that is, the restructuring of corporate resources, routines, and operating patterns—is perhaps the most sensitive to the functioning of the family system. Change in family businesses is often not just an organizational or strategic issue, but also a reinterpretation of family identity. A new business model, a leadership transition, a shift toward digitalization, or internationalization can easily raise the question of what will remain of the family legacy. The family council can support the capacity for transformation if it interprets change not as a rejection of the past, but as a means of ensuring the family business's continuity. If, however, the council equates the protection of tradition with the rejection of change, then the capacity for transformation may weaken.

3.2 Mediating and moderating variables in the FC–DC relationship

The relationship between family council and dynamic capabilities does not operate directly, automatically, or in the same way in every family business. Based on a comparison of the literature, several variables can be identified that may mediate or moderate this relationship. These are not elements of the same nature. Some function more as processes, others as relational resources, and still others as contextual conditions. Therefore, it is not appropriate to uniformly refer to all of them as "mechanisms." In the model, we treat them as mediating and moderating variables.

Balancing the aspects related to social-emotional capital is one of the most important mediating processes. In family businesses, strategic decisions often have not only economic but also social-emotional consequences. Fear of losing control, protecting the family name, attachment to the founder's legacy, or the intention to pass the business on to the next generation can all influence what kind of adaptation the family considers acceptable. The family council can support dynamic capabilities if it does not suppress these considerations but also does not allow them to completely dominate strategic decisions. The challenge is to strike a balance between family dynamics and the company's adaptation needs.

Family social capital is also important. Trust, a shared language, the density of relationships, and the willingness of family members to cooperate can reduce the cost of decision-making and accelerate the mobilization of resources. At the same time, overly closed social capital can restrict the flow of information and reinforce the echo chamber effect. Therefore, family social capital is not a positive factor in and of itself. The question is whether it supports the acceptance of differing information, constructive debate, and strategic change.

Intergenerational knowledge transfer is particularly important from the perspectives of sensing and transforming. In family businesses, significant knowledge can accumulate in tacit form: in relationships, stories, decision-making patterns, corporate memory, and industry experience. The family council can help make this knowledge visible and transferable. At the same time, intergenerational knowledge transfer cannot simply mean adopting the older generation's way of thinking. In terms of dynamic capabilities, it becomes valuable only if old knowledge can be reinterpreted in new situations.

Legitimacy-building and conflict resolution also influence the FC–DC relationship. Strategic adaptation often involves losses, uncertainty, and role changes. If the family council is able to legitimize the change, the company can more easily move from recognition to action. If conflicts are avoided or suppressed, decision-making paralysis may persist behind the facade of apparent agreement.

Several moderating factors can shape this relationship. The generational stage, family complexity, concentration or dispersion of ownership, the maturity of the family council, the size of the company, industry dynamics, and the cultural context can all influence the extent to which the family council's characteristics and functioning impact dynamic capabilities. In a first-generation, founder-led company, the role of the family council may differ from that in a multi-branch, multi-generational family of owners. In a stable, slowly changing industry, a different type of council operation may suffice than in the face of strong technological or market turbulence.

3.3 Dynamic capabilities and firm Performance

The third element of the model is the relationship between dynamic capabilities and firm performance. According to the theory of dynamic capabilities, the ultimate significance of sensing, seizing, and transforming lies in the firm's ability to adapt, renew itself, and maintain its competitiveness in a changing environment (Teece, 2007). However, the impact on performance is not self-evident. Not every instance of sensing leads to action, not every strategic decision leads to successful implementation, and not every transformation improves the company's performance.

Particular caution is required when interpreting the DC–performance relationship. The literature often identifies mediating variables that are conceptually close to dynamic capabilities themselves. Innovation, for example, can be an output of dynamic capabilities, but in many studies it appears partly as a component of seizing or transforming. In many cases, knowledge integration is more of a micro-condition of dynamic capabilities rather than a clear mediating variable between them and performance. Strategic flexibility is also difficult to separate from the functions of sensing and transforming. Therefore, the present model does not base its explanation of the DC–performance relationship on these.

In the model, dynamic capabilities contribute to corporate performance when sensing, resource mobilization, and transformation are transformed into coordinated, timely, and resource-efficient corporate action. Here, corporate performance primarily refers to operational efficiency, market responsiveness, financial performance, speed of adaptation, and resilience. The emphasis, therefore, is not on the fact that dynamic

capabilities are generally “good” for performance, but rather on the conditions under which they can actually be translated into performance.

3.4 The family council as a moderator of the DC–corporate performance relationship

A central tenet of the model is that the family council influences not only the development of dynamic capabilities but also their impact on performance. Therefore, the characteristics and functioning of the family council can also be interpreted as a moderator of the DC–corporate performance relationship. This does not mean that the family council mediates between dynamic capabilities and performance. The family council does not follow dynamic capabilities in the causal chain. Rather, it influences whether the established sensing, seizing, and transforming processes receive family support, owner legitimacy, and sustained resource commitment.

This moderating role is particularly important in the case of seizing and transforming. A company may perceive a market opportunity, but if the family council is unable to build owner consensus, seizing the opportunity may stall. The company may be capable of organizational restructuring, but if the family interprets the change as a threat to family identity, the restructuring may provoke resistance. The reverse is also true: a well-functioning family council can lend legitimacy to change, help the family interpret risks, and provide long-term support for decisions that are uncertain in the short term but strategically necessary.

The family council thus influences the DC–performance relationship by reducing or increasing organizational friction regarding dynamic capabilities. If the council functions well, adaptive capabilities can be more easily translated into coordinated corporate action. If it functions poorly, those same capabilities may remain partially untapped or be lost amid family conflicts, decision-making delays, and a lack of legitimacy.

3.5 Recursive feedback loops

The model does not stop at the relationship between the DC and corporate performance. It also assumes two feedback loops. The first feeds back to the family council from the dynamic capabilities. As the company develops its sensing, seizing, and transforming capabilities, new types of governance needs may emerge. The family council must manage different information, involve new stakeholders, legitimize more complex decisions, and, over time, adapt its own operations as well. The development of dynamic capabilities therefore requires not only corporate but also family governance learning.

The second feedback loop comes from corporate performance. Successful performance can strengthen the family council’s legitimacy, increase the family’s confidence in joint decision-making processes, and reinforce the council’s strategic role. Poor performance, on the other hand, can trigger conflicts, blame, loss of trust, or demands for governance reform. Performance is thus not merely the endpoint of the model but also a form of feedback that shapes the characteristics and functioning of the family council.

For this reason, the model is fundamentally recursive. The family council shapes dynamic capabilities; dynamic capabilities influence corporate performance; and the family council moderates this relationship. Meanwhile, dynamic capabilities and corporate performance feed back into the family council. Thus, the adaptation of a family business is not a one-time causal chain, but a continuous governance–capability–performance cycle.

3.6 Propositions

P1. The characteristics and functioning of the family council influence the development of dynamic capabilities in family businesses by shaping the quality of the sensing, seizing, and transforming processes.

P2. The relationship between the characteristics and functioning of the family council and dynamic capabilities may be mediated by family and governance variables such as the balancing of social-emotional aspects, family social capital, intergenerational knowledge transfer, legitimacy building, conflict resolution, and the promotion of ambidexterity.

P3. The strength and direction of the relationship between the characteristics and functioning of the family council and dynamic capabilities are moderated by family, ownership, governance, and external contextual conditions, such as generational stage, family complexity, ownership structure, council maturity, industry dynamism, and cultural context.

P4. Dynamic capabilities contribute to corporate performance when the sensing, seizing, and transforming processes are translated into timely, coordinated, and resource-efficient corporate action.

P5. The characteristics and functioning of the family council moderate the relationship between dynamic capabilities and corporate performance by influencing family support for dynamic capabilities, their ownership legitimacy, resource availability, and organizational acceptance.

P6. The development of dynamic capabilities has an impact on the family council because it creates new governance demands, learning tasks, coordination needs, and operational expectations.

P7. Corporate performance affects the family council by shaping its legitimacy, perceived usefulness, level of conflict, strategic priorities, and willingness to undergo management renewal.

4 Discussion

The most important theoretical claim of the proposed model is that the family council should not be viewed merely as a forum for family communication or conflict resolution. The family council is one possible governance infrastructure for the adaptation of the family business. Its impact does not stem from its formal existence, but rather from its characteristics, how it operates, and its ability to organize family intentions, tensions, and preferences into a form that can be processed for corporate adaptation. In doing so, the model clarifies the relationship between the literature on family governance and that on dynamic capabilities. The family council is not simply a "family background factor," but a governance space that can influence what the company perceives, what opportunities it seizes, and what transformations it considers legitimate.

The model's second contribution is the clarification of the family council's dual role. On the one hand, the family council is a precursor to dynamic capabilities: its characteristics and functioning influence the development of sensing, seizing, and transforming processes. On the other hand, the family council also acts as a moderator of the relationship between dynamic capabilities and firm performance. This means that the performance impact of dynamic capabilities may depend on whether the family is able to provide support, ownership legitimacy, resource commitment, and acceptance for adaptive action. The model therefore does not position the family council as a mediator between dynamic capabilities and performance. Instead, it posits that the council is simultaneously involved in the development of capabilities and in the conditions for translating them into performance.

From a practical perspective, this means that when evaluating family counseling, it is not sufficient to examine whether it improves family communication or reduces conflicts. These are important outcomes, but they do not capture the strategic significance of counseling. It must also be examined whether the council contributes to the perception of environmental changes, the seizing of strategic opportunities, the reallocation of resources, and the family's legitimization of changes. A family council truly becomes a strategic forum when it not only maintains family unity but also helps the company adapt more effectively.

The model also identifies several directions for future research. First, the FC–DC relationship must be examined more closely: which family council characteristics and operational features have the strongest impact on the sensing, seizing, and transforming dimensions. Second, it must be clarified when the identified variables—such as the balancing of socio-emotional aspects, family social capital, intergenerational knowledge transfer, legitimacy, and conflict resolution—function as mediators, moderators, or contextual conditions. Third, it remains unclear exactly through which processes dynamic capabilities translate into firm performance; this requires particularly careful examination, as several previously used mediators conceptually overlap with the dynamic capabilities themselves. Fourth, it must be empirically tested whether the characteristics and functioning of the family council truly moderate the DC–performance relationship. Finally, the external context requires more attention: we do not know exactly how market dynamism, the intensity of competition, technological turbulence, or industry maturity interact with the functioning of the family council. Longitudinal studies are also needed to understand how dynamic capabilities and corporate performance influence the legitimacy, composition, and functioning of the family council.

5 Conclusion

This study suggests that the strategic significance of the family council should not be understood merely in terms of its existence, but rather in relation to its dynamic capabilities and corporate performance. The family council can become an important institution for the family business's adaptation, but only if it can channel the family system's values, tensions, and long-term preferences to support the processes of sensing, seizing, and transforming. The proposed model is therefore not a normative statement about the usefulness of the family council, but rather a framework for future empirical research to better understand the relationship between family governance, dynamic capabilities, and corporate performance.

6 Ethics Declaration

This paper is based on a structured integrative literature synthesis and does not involve human participants, personal data, interviews, experiments, or fieldwork. Therefore, ethical clearance was not required.

7 AI Declaration

The authors declare that no AI tools were applied to develop the conference paper.

References

Chrisman, J.J., Chua, J.H., Le Breton-Miller, I., Miller, D. and Steier, L.P. (2018) 'Governance mechanisms and family firms', Entrepreneurship Theory and Practice, 42(2), pp. 171–186.

Eisenhardt, K.M. and Martin, J.A. (2000) 'Dynamic capabilities: What are they?', Strategic Management Journal, 21(10–11), pp. 1105–1121.

Gersick, K.E., Davis, J.A., McCollom Hampton, M. and Lansberg, I. (1997) Generation to Generation: Life Cycles of the Family Business. Boston, MA: Harvard Business School Press.

Gómez-Mejía, L.R., Haynes, K.T., Núñez-Nickel, M., Jacobson, K.J.L. and Moyano-Fuentes, J. (2007) 'Socioemotional wealth and business risks in family-controlled firms: Evidence from Spanish olive oil mills', Administrative Science Quarterly, 52(1), pp. 106–137.

Habbershon, T.G. and Williams, M.L. (1999) 'A resource-based framework for assessing the strategic advantages of family firms', Family Business Review, 12(1), pp. 1–25.

Snyder, H. (2019) 'Literature review as a research methodology: An overview and guidelines', Journal of Business Research, 104, pp. 333–339.

Suess, J. (2014) 'Family governance: Literature review and the development of a conceptual model', Journal of Family Business Strategy, 5(2), pp. 138–155.

Teece, D.J. (2007) 'Explicating dynamic capabilities: The nature and microfoundations of sustainable enterprise performance', Strategic Management Journal, 28(13), pp. 1319–1350.

Teece, D.J., Pisano, G. and Shuen, A. (1997) 'Dynamic capabilities and strategic management', Strategic Management Journal, 18(7), pp. 509–533.

Webster, J. and Watson, R.T. (2002) 'Analyzing the past to prepare for the future: Writing a literature review', MIS Quarterly, 26(2), pp. xiii–xxiii.

Analytics Maturity and AI Governance for Organizational Trust: A Swiss Capability Perspective

Viktoriia Apalkova
University of Applied Sciences and Arts Western Switzerland, Fribourg, Switzerland

Abstract

This study examines how analytics maturity and AI governance are articulated as organizational capabilities that support trust in AI-driven and data-based decision-making. As organizations increasingly rely on artificial intelligence (AI) and advanced analytics, trust has become a critical concern, extending beyond technical performance to include reliability, responsibility, and institutional alignment. While existing research often treats analytics maturity as a linear progression and AI governance as a set of external constraints, this study adopts a configurational and capability-based perspective grounded in the Resource-Based View (RBV) and Dynamic Capabilities Theory (DCT).

Drawing on qualitative analysis of corporate website discourse from ten Swiss organizations across multiple industries, the study explores how AI applications, data infrastructures, governance mechanisms, and trust-related narratives are publicly communicated. The findings demonstrate that analytics maturity is not represented as a linear developmental path but as distinct configurations of interdependent resources, combining technical, organizational, and governance capabilities. AI governance emerges as an embedded organizational capability that supports trust by enabling reliability, accountability, and alignment with institutional expectations, rather than functioning solely as a compliance mechanism.

Trust is constructed differently across sectors, reflecting industry-specific logics such as clinical validation in healthcare, risk control in finance, data sovereignty in telecommunications, and operational safety in industrial and energy contexts. These findings highlight that trust in AI is not universal but context-dependent and institutionally embedded.

The study contributes by reconceptualizing analytics maturity as a configurational phenomenon, repositioning AI governance as an enabling capability, and demonstrating how trust is co-constructed through the alignment of organizational resources and institutional expectations. The results underline the importance for organizations to align technical capabilities with governance structures and communication practices to sustain trust in AI-enabled environments..

Keywords: Analytics maturity, AI governance, Organizational trust, Dynamic capabilities, Switzerland.

Paper type: Academic Research Paper

1 Introduction

The increasing reliance on artificial intelligence (AI) and advanced analytics in organizational decision-making has brought ***organizational trust*** to the forefront of both academic and managerial debates. As AI systems increasingly influence critical decisions, questions arise not only about their technical performance, but also about whether organizations can be trusted to design, govern, and use AI responsibly. Recent research suggests that trust in AI-enabled decision-making depends less on algorithmic sophistication and more on organizational capabilities related to analytics maturity and AI governance (Bakonyi, Z., 2024 ; Smit, D. et al., 2021 ; Lahusen, C. et al., 2024).

Organizational analytics maturity refers to the extent to which data and analytics are systematically embedded in organizational structures, routines, and decision-making processes (Grossman, R. L., 2018; Lismont, J. et al., 2017). High levels of analytics maturity are associated with coordinated data practices, shared analytical understanding, and the ability to integrate analytics into strategic and operational decisions (Król, K. et al., 2020). However, much of the existing analytics maturity literature remains technocentric, emphasizing data infrastructure and analytical tools while paying limited attention to how maturity contributes to trust in ***data-driven decisions***, both internally and externally (Ariyarathna, K. et al., 2019; Langer, B., 2025; Steinmetz, S. et al., 2025).

In parallel, the literature on ***AI governance*** has expanded rapidly, focusing on ethical principles, regulatory compliance, and technical safeguards such as transparency, explainability, and fairness (Bahangulu, J. K. et al., 2025; Birkstedt, T. et al. 2023; Choung, H. et al., 2024). While these studies provide essential normative guidance, AI governance is often conceptualized as a set of constraints imposed on organizations (Batool, A. et al., 2025), rather than as an enabling organizational capability that supports trust in AI systems and analytics-based decision-making.

From a ***Resource-Based View*** (RBV), analytics maturity and AI governance can be understood as intangible, knowledge-based organizational resources embedded in routines, governance structures, and shared interpretations of data and AI use (Moderno, O. B. D. S. et al. , 2024). These resources are socially complex and

difficult to imitate, making them central to the development of organizational trust. From a ***Dynamic Capabilities Theory*** (DCT) perspective, analytics maturity and AI governance enable organizations to continuously align AI practices with evolving technological, regulatory, and societal expectations, thereby sustaining trust over time rather than achieving short-term performance gains (Lahusen, C. et al., 2024; Herrera-Poyatos, A. et al., 2025).

The ***Swiss organizational context*** offers a particularly relevant setting for examining these dynamics. Switzerland is characterized by high institutional trust, strong regulatory frameworks, and a knowledge-intensive economy in which reliability, accountability (Albu, O. B. et al., 2019), and transparency are central organizational values. In such an environment, trust in AI and data-driven decision-making is not assumed but must be actively constructed and maintained through visible governance practices and mature analytics capabilities.

Despite the growing relevance of trust in AI, existing research rarely examines how analytics maturity and AI governance jointly function as organizational capabilities that support trust, nor how organizations publicly articulate these capabilities. Moreover, empirical studies often rely on surveys or interviews, providing limited insight into how trust is built and communicated through organizational digital discourse. This study addresses this gap by adopting a capability-based perspective on analytics maturity and AI governance and by examining how Swiss organizations articulate trust-building practices in relation to AI and analytics.

Grounded in RBV and DCT, the study is guided by the following research questions focused on organizational trust in AI and data-driven decision-making:

- ***RQ1:*** *How do Swiss organizations publicly articulate analytics maturity and AI governance as foundations for organizational trust in AI and data-driven decision-making?*
- ***RQ2:*** *What governance-related knowledge resources and organizational structures are emphasized to support trust in the use of AI and analytics?*
- ***RQ3:*** *How do different configurations of analytics maturity and AI governance reflect distinct approaches to building and maintaining organizational trust in the Swiss context?*

2 Methodology

This study adopts a qualitative, multiple-case research design to examine how organizations articulate analytics maturity, AI governance, and organizational trust through publicly available corporate communication. The objective is not to assess internal technological performance, but to analyze how AI application is constructed, described, and legitimized in organizational discourse, with a particular focus on the relationship between analytics maturity and trust. A qualitative approach is appropriate given the exploratory nature of the research and its emphasis on meaning-making processes and semantic interpretation.

The study is based on data collected from official corporate websites, including innovation and AI-related pages, governance and sustainability sections, media releases, and career-related materials. Corporate websites are treated as institutional communication artifacts, reflecting both organizational practices and deliberate efforts to present capabilities, responsibility, and strategic direction to external stakeholders. Website content represents a form of unstructured textual data, which can be systematically analyzed to extract insights about organizational behavior and positioning. This approach aligns with recent methodological developments that emphasize the analytical value of unstructured web data and its interpretation through qualitative and natural language-based techniques (Pichiyan et al., 2023).

A purposive sampling strategy was used to select ten organizations operating in Switzerland, ensuring diversity across industries, regulatory environments, and AI application contexts. The sample includes companies from healthcare, finance, insurance, telecommunications, energy, industrial technology, consumer goods, and mining. This diversity allows capturing variation in how AI is applied, governed, and communicated, supporting a configurational perspective on analytics maturity and trust.

Table 1: Sample of analyzed companies

Company	Industry	Key characteristic*
Nestlé	Food & Beverage	AI across value chain and consumer services
ABB	Industrial Technology	Industrial AI platforms and analytics
Novartis	Pharmaceuticals	AI-driven R&D and innovation ecosystem
Roche	Healthcare	AI in diagnostics and clinical processes
UBS	Banking	AI for risk management and financial services
Zurich Insurance Group	Insurance	AI in underwriting and claims
Swiss Re	Reinsurance	AI for risk modelling and analytics
Swisscom	Telecommunications	AI platforms and data sovereignty
Axpo Group	Energy	AI in critical infrastructure operations

Company	Industry	Key characteristic*
Glencore	Mining	Automation and safety-driven AI applications

*Company website sources:
Nestlé. https://www.nestle.com/about/research-development/digital-innovation
ABB. https://global.abb/group/en/technology/did-you-know/abb-ability-genix-industrial-analytics-and-ai-suite
Novartis. https://www.novartis.com/stories/building-future-novartis-innovation-lab-ai-game-studio-xr-emerging-tech
Roche. https://www.roche.com/stories/ai-revolutionising-drug-discovery-and-transforming-patient-care
UBS. https://www.ubs.com/global/en/our-firm/what-we-do/technology/innovation-and-ai.html
Zurich Insurance Group. https://www.zurich.com/about-us/ai-at-zurich
Swiss Re. https://www.swissre.com/risk-knowledge/advancing-societal-benefits-digitalisation.html
Swisscom. https://www.swisscom.ch/en/business/enterprise/offer/platforms-applications/data-driven-business/swiss-ai-platform.html
Axpo Group. https://www.axpo.com/ch/en/energy/digital-solutions/digital-transformation-nuclear/ai-assistant.html
Glencore. https://www.glencore.com/media-and-insights/insights/using-advanced-autonomous-technology-to-keep-our-people-safe

The analysis follows a qualitative coding approach focused on the semantic interpretation of AI application discourse. Rather than identifying only the presence of AI technologies, the study examines how organizations describe AI, what meanings are attached to it, and how it is linked to broader constructs such as strategy, governance, and trust. This approach is grounded in interpretive analysis of textual data, where language is understood as a mechanism through which organizations construct and communicate their capabilities and intentions.

The coding framework was developed using a hybrid approach, combining concept-driven categories from existing literature with iterative refinement during analysis. In particular, the study conceptualizes trust as a latent construct that is not directly measurable in website data but can be inferred from organizational discourse on reliability, responsibility, and stakeholder impact (Patent & Searle, 2019). Based on this perspective, the following coding categories were applied: AI_USE (descriptions of AI applications), DATA_MATURITY (data and analytics capabilities), STRATEGY (strategic positioning of AI), GOVERNANCE (rules, controls, and compliance mechanisms), ETHICS (fairness, transparency, responsibility), TRUST (explicit or implicit trust-related narratives), and CAPABILITY (organizational skills, structures, and learning processes).

Coding used short textual segments to capture precise semantic expressions and ensure consistency. The analysis was iterative, combining within-case interpretation with cross-case comparison to identify patterns in how AI application is constructed and related to governance and trust across institutional contexts.

An institutional perspective informs the interpretation, recognizing that corporate communication is shaped by regulation, industry norms, and legitimacy pressures. AI discourse is therefore understood not only as reflecting technological capability but also as a strategic communication practice through which organizations signal alignment with societal expectations and build trust. Differences across companies are interpreted as reflecting institutional variation rather than purely technological differences.

Analytical rigor is ensured through systematic coding, multiple-case comparison, and the use of official sources. The reliance on publicly available data may emphasize strategic representation over internal practices, but it provides valuable insights into how organizations construct and communicate AI capabilities, governance, and trust.

3 Results

The analysis reveals that the articulation of AI application across Swiss organizations cannot be reduced to a single dimension of technological adoption or analytics maturity. Rather than conceptualizing analytics maturity as a linear progression, this study interprets it through distinct organizational configurations, where technical capabilities, organizational structures, governance arrangements, and trust-oriented narratives are combined in different ways. This finding challenges traditional maturity models and supports a configurational understanding of AI as a bundle of interdependent organizational resources embedded in institutional contexts.

At the same time, trust emerges as a central organizing concept across all cases. Trust is interpreted as a context-dependent construct encompassing perceptions of reliability, responsibility, and institutional alignment. This interpretation allows capturing the diverse ways in which organizations construct trust in relation to AI and analytics, reflecting differences in industry requirements, regulatory pressures, and stakeholder expectations.

At the most immediate level, AI is consistently presented as an operational and technological resource, embedded in core processes such as drug discovery, financial risk management, claims processing, industrial optimization, and infrastructure management. Organizations describe AI through concrete applications, such as “AI… identify molecules” (Novartis), “detect fraud and anticipate market trends” (UBS), or “automate claims handling” (Swiss Re). Across cases, these applications are closely linked to data infrastructures, with repeated emphasis on “large volumes of unstructured data” (Swiss Re), “predictive analytics” (Nestlé), and “data-driven

business" (Swisscom). This indicates that AI is not conceptualized as a standalone technology but as a data-dependent analytical resource, forming the technical foundation of analytics maturity.

This technical resource base is embedded within an organizational layer, where AI is framed as a strategic and capability-building resource. Companies explicitly position AI as central to their long-term development, as reflected in formulations such as "AI360 strategy" (Zurich), "AI-enabled institution" (UBS), and "transform patient care" (Roche). In parallel, organizations emphasize capability development through AI labs, training programs, interdisciplinary expertise, and partnerships. These elements indicate that analytics maturity is constructed as an organizational capability, reflecting the ability to integrate, deploy, and scale AI-related resources across the firm.

Beyond these dimensions, the findings highlight a governance layer, where AI governance is articulated as a complementary resource ensuring the responsible and controlled use of AI. In highly regulated sectors such as healthcare, finance, and insurance, governance is explicitly formalized through compliance frameworks, validation procedures, and ethical principles. For example, Novartis refers to an "AI Risk & Compliance Management Framework," while Roche emphasizes that AI "must undergo rigorous clinical validation." Similarly, financial institutions highlight "strong oversight" (UBS) and regulatory alignment (Swiss Re). In contrast, organizations in less regulated sectors, such as consumer goods (Nestlé) and industrial technology (ABB), provide only limited or implicit references to AI governance. In the case of Glencore, governance is present but framed through safety systems rather than AI-specific mechanisms. These differences indicate that governance resources are selectively developed and communicated, depending on institutional context and stakeholder expectations.

At a higher level, the analysis reveals an institutional layer, where AI-related resources are linked to trust, legitimacy, and societal alignment. Across all cases, AI is associated with trust-related outcomes, yet the mechanisms through which trust is constructed differ significantly. In healthcare, trust is grounded in clinical validation and patient outcomes ("safe and reliable healthcare-grade AI" – Roche), emphasizing reliability and responsibility. In finance, trust is associated with risk control and responsible decision-making ("responsible and sustainable AI" – UBS), reflecting institutional expectations of accountability. In telecommunications, trust is framed through data sovereignty and security ("secure, trustworthy AI solutions made in Switzerland" – Swisscom), highlighting institutional alignment with national data standards. In energy and mining, trust is linked to safety and operational reliability ("reliable operation of plants" – Axpo; "everyone should return home unharmed" – Glencore), emphasizing responsibility toward infrastructure and workforce. These findings demonstrate that trust is not a universal outcome of AI adoption, but an institutionally embedded construct shaped by industry-specific logics.

Table 2: Cross-case comparison of AI application and governance articulation

Company	AI application focus	Governance articulation	Trust construction
Novartis	R&D, innovation	Strong (formal frameworks)	Patient outcomes
Roche	Clinical AI, diagnostics	Very strong (validation)	Safety & reliability
UBS	Risk, advisory	Strong (oversight)	Risk control
Zurich	Claims, underwriting	Strong (training-based)	Customer resilience
Swiss Re	Risk modelling	Strong (compliance)	Data transparency
Swisscom	AI platforms	Strong (data security)	Data sovereignty
Axpo	Infrastructure operations	Strong (regulatory)	Safety & reliability
ABB	Industrial analytics	Limited (implicit)	Performance
Nestlé	Consumer AI	Limited (implicit)	Product/service trust
Glencore	Automation & safety	Strong (non-AI governance)	Worker safety

The comparative analysis further reveals that organizations cluster into distinct configurations of AI-related resources, reflecting different approaches to building and maintaining trust.

Table 3: Configurations of AI-related resources and trust

Configuration	Companies	Resource characteristics	Trust logic
Integrated resource configuration	Nestlé, ABB	Strong AI and data resources, weak governance articulation	Implicit (performance-based reliability)
Trust-oriented resource configuration	Novartis, Roche, UBS, Zurich, Swiss Re, Swisscom, Axpo	Strong alignment of AI, governance, and ethical resources	Explicit (regulated and institutional trust)
Operational–safety resource configuration	Glencore	AI embedded in safety and operational control systems	Operational trust (safety and responsibility)

These configurations demonstrate that organizations do not converge toward a single model of analytics maturity. Instead, they exhibit qualitatively distinct combinations of AI capability, governance, and institutional alignment. Figure 1 summarizes the multi-layered configuration of AI-related resources identified across cases.

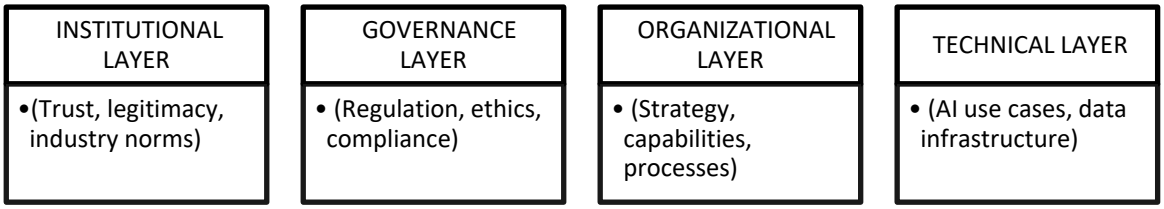


Figure 1: T. Conceptual framework: AI application as a multi-layered construct

The empirical findings provide clear answers to the research questions. First, Swiss organizations articulate analytics maturity as the development and integration of AI-related technical and organizational resources, while governance is positioned as a complementary resource ensuring responsible and controlled use. Trust, understood as a context-dependent construct encompassing reliability, responsibility, and institutional alignment, is constructed as the outcome of this resource configuration rather than as an independent objective. Second, governance-related knowledge resources include formal frameworks, validation mechanisms, compliance structures, and organizational capabilities such as training and interdisciplinary expertise, all of which contribute to the credibility and reliability of AI systems. Third, different configurations of analytics maturity and AI governance reflect distinct, institutionally shaped approaches to trust construction, depending on industry context and regulatory environment.

4 Discussion

The findings challenge dominant assumptions in the analytics maturity literature by demonstrating that maturity is not best understood as a linear progression of technological sophistication. Instead, it is more accurately conceptualized as a configuration of interdependent organizational resources, where value emerges from alignment rather than accumulation. This shifts attention away from staged development models toward the structural integration of capabilities, highlighting that similar levels of technological advancement can produce different outcomes depending on how resources are combined.

AI governance, typically framed as a set of external constraints or normative requirements, appears here as an organizationally embedded resource with both operational and symbolic functions. Governance mechanisms do not merely regulate AI use but actively contribute to the construction of reliability and responsibility. Their role extends beyond compliance toward enabling organizations to demonstrate alignment with institutional expectations, thereby reinforcing trust. This dual function suggests that governance should be understood as part of the organization's capability architecture, rather than as an external overlay.

Trust itself does not emerge from technological performance alone but from the coherence between resources and institutional context. The variation in trust logics across industries indicates that trust is constructed through different mechanisms—validation in healthcare, risk control in finance, data sovereignty in telecommunications, and safety in industrial environments. These differences reinforce the interpretation of trust as context-dependent, shaped by sectoral norms and expectations rather than universally defined criteria.

From a Resource-Based View perspective, analytics maturity and AI governance function as intangible, socially complex resources that extend beyond efficiency gains. Their strategic relevance lies in their contribution to organizational legitimacy and trustworthiness, particularly in environments characterized by regulatory intensity and societal scrutiny. This expands the RBV by emphasizing non-economic value creation, where trust becomes a critical outcome of resource configuration.

Dynamic Capabilities Theory further clarifies these differences. The observed configurations reflect varying abilities to align and reconfigure resources over time, particularly in response to evolving institutional demands. Configurations characterized by strong alignment between AI capabilities and governance structures indicate more advanced dynamic capabilities, where organizations continuously adapt not only their technologies but also their governance practices and communication strategies. Trust, in this sense, becomes a dynamic outcome, sustained through ongoing capability development rather than achieved through static implementation.

The Swiss context reinforces these dynamics. The prominence of governance and trust-oriented configurations reflects broader institutional conditions in which reliability, accountability, and transparency are central expectations. Organizational practices and their articulation are therefore shaped not only by internal

capabilities but also by institutional pressures that define what constitutes trustworthy AI. This highlights the importance of considering national and sectoral contexts when analyzing AI governance and trust.

5 Conclusions

A configurational perspective offers a more appropriate lens for understanding analytics maturity and AI governance. Rather than progressing through predefined stages, organizations develop distinct capability combinations shaped by internal resource alignment and external institutional expectations. Within these configurations, trust does not result directly from technological advancement but from how capabilities are structured, governed, and communicated.

Analytics maturity and AI governance function as complementary capabilities that jointly shape perceptions of reliability and responsibility. Their interdependence suggests that technical improvements alone are insufficient for sustaining trust, particularly where accountability and transparency are critical.

Three contributions follow. First, analytics maturity is reconceptualized as a configurational phenomenon, shifting the focus from progression to alignment. Second, AI governance is repositioned as an enabling capability integral to trust construction. Third, trust in AI is shown to be institutionally embedded, varying across sectors and shaped by context-specific expectations.

Practically, organizations should prioritize alignment of technical, organizational, and governance capabilities rather than treating them as separate initiatives. Trust emerges from coherence, implying that AI investments must be accompanied by parallel developments in governance, capabilities, and communication.

The reliance on website data limits insight into internal practices, suggesting future research should combine discourse analysis with in-depth studies. Extending analysis across national contexts would further clarify institutional influences. Sustaining trust in AI ultimately depends less on technological advancement alone than on continuous alignment with evolving expectations.

6 Ethics Declaration

This study is based exclusively on publicly available data collected from official corporate websites. No human participants were involved, and no personal or sensitive data were collected or processed. The research complies with applicable ethical standards for academic research and does not require formal ethical approval.

7 AI Declaration

The author declares that artificial intelligence tools were used solely for language refinement and editing support. All conceptual development, data collection, analysis, and interpretation were conducted independently by the author. The author takes full responsibility for the content and integrity of the work.

References

Albu, O. B., & Flyverbom, M. (2019). Organizational transparency: Conceptualizations, conditions, and consequences. Business & society, 58(2), 268-297.

Ariyarathna, K., & Peter, S. (2019). Business analytics maturity models: a systematic review of literature. Focus, 3(10), 4.

Bahangulu, J. K., & Owusu-Berko, L. (2025). Algorithmic bias, data ethics, and governance: Ensuring fairness, transparency and compliance in AI-powered business analytics applications. World J Adv Res Rev, 25(2), 1746-63.

Bakonyi, Z. (2024). How can companies handle paradoxes to enhance trust in artificial intelligence solutions? A qualitative research. Journal of Organizational Change Management, 37(7), 1405-1426

Batool, A., Zowghi, D., & Bano, M. (2025). AI governance: a systematic literature review. AI and Ethics, 1-15.

Birkstedt, T., Minkkinen, M., Tandon, A., & Mäntymäki, M. (2023). AI governance: themes, knowledge gaps and future agendas. Internet Research, 33(7), 133-167.

Choung, H., David, P., & Seberger, J. S. (2024). A multilevel framework for AI governance. In The Routledge Handbook of Global and Digital Governance Crossroads (pp. 310-323). Routledge India.

Grossman, R. L. (2018). A framework for evaluating the analytic maturity of an organization. International Journal of Information Management, 38(1), 45-51.

Herrera-Poyatos, A., Del Ser, J., López de Prado, M., Wang, F. Y., Herrera-Viedma, E., & Herrera, F. (2025). Responsible Artificial Intelligence Systems: A Roadmap to Society's Trust through Trustworthy AI, Auditability, Accountability, and Governance. arXiv e-prints, arXiv-2503.

Król, K., & Zdonek, D. (2020). Analytics maturity models: An overview. Information, 11(3), 142.

Lahusen, C., Maggetti, M., & Slavkovik, M. (2024). Trust, trustworthiness and AI governance. Scientific Reports, 14(1), 20752.

Langer, B. (2025). Understanding Data & Analytics Maturity: A Systematic Review of Maturity Model Composition. Schmalenbach Journal of Business Research, 1-23.

Lismont, J., Vanthienen, J., Baesens, B., & Lemahieu, W. (2017). Defining analytics maturity indicators: A survey approach. International Journal of Information Management, 37(3), 114-124.

Moderno, O. B. D. S., Braz, A. C., & Nascimento, P. T. D. S. (2024). Robotic process automation and artificial intelligence capabilities driving digital strategy: a resource-based view. Business Process Management Journal, 30(1), 105-134.

Patent, V., & Searle, R. H. (2019). Qualitative meta-analysis of propensity to trust measurement. Journal of Trust Research, 9(2), 136-163.

Pichiyan, V., Muthulingam, S., Nalajala, S., Ch, A., & Das, M. N. (2023). Web scraping using natural language processing: exploiting unstructured text for data extraction and analysis. Procedia Computer Science, 230, 193-202.

Smit, D., Eybers, S., & Smith, J. (2021). A data analytics organisation's perspective on trust and AI adoption. In Southern African Conference for Artificial Intelligence Research (pp. 47-60). Cham: Springer International Publishing.

Steinmetz, S. T., Naugle, A., Schutte, P., Sweitzer, M., Washburne, A., Linville, L., ... & Myren, S. (2025). The Trust Calibration Maturity Model for Characterizing and Communicating Trustworthiness of AI Systems. arXiv preprint arXiv:2503.15511.

Redefining Financial Advisory: Human-Machine Collaboration and Organisational Transformation Under the EU Financial Data Access Regulation

Gabriele Rizzo[1], Daniele Allega[12]
[1] LUISS Guido Carli University, Rome, Italy
[2] Mercatorum University, Rome, Italy

Abstract

The convergence of artificial intelligence and the proposed European Financial Data Access Regulation (FiDA) is reshaping financial advisory in ways that affect both client relationships and the structural composition of financial institutions. The dominant discourse has long oscillated between two extreme narratives: technological utopianism, which forecasts the displacement of human advisors, and technological scepticism, which confines algorithms to instrumental roles. This paper articulates a third narrative, that of human-machine symbiosis, in which AI capabilities and human expertise are deliberately coordinated as complementary contributors to advisory decision-making, and shows that this view carries macro-financial consequences that have received insufficient attention. The argument unfolds in four steps. We first review four interconnected literatures: AI in financial advisory, human-machine collaboration, FiDA and Open Finance, and capital structure theory in banking. We then develop a conceptual framework in which client data, algorithmic insight, and advisor expertise are coordinated as knowledge assets, deriving two implementation scenarios that distinguish AI-as-empowerment from AI-as-autonomous-interface configurations. A design-science proof-of-concept operationalises the framework on FiDA-aligned synthetic data, evaluated against alignment, suitability-augmentation, and interpretability criteria. Finally, we develop the macro-financial implications. AI-mediated advisory under FiDA can serve as a demand-side mechanism that complements the supply-side capital pressures of Basel III/IV. Through three transmission channels, ranging from direct conversion of deposits into Tier 1 or Tier 2 capital to maturity-profile improvements, depositors are progressively transformed into stakeholders, lowering institutional leverage and enhancing resilience to adverse macroeconomic shocks. The contribution is fourfold: a theoretically grounded alternative to dominant narratives; a mapping of FiDA provisions onto AI-advisory operational requirements; a development of macro-financial implications linking micro-personalisation to capital structure dynamics; and an algorithm-agnostic proof-of-concept.

Keywords: Human-Machine Collaboration, FiDA, Artificial Intelligence, Financial Advisory, Capital Structure

Paper type: Academic Research Paper

1 Introduction

The financial services industry is undergoing a profound transformation driven by artificial intelligence (AI), with implications extending far beyond operational efficiency to the structural composition of financial institutions (Agrawal et al., 2018; D'Acunto and Rossi, 2023). Yet the dominant discourse has long oscillated between two extremes. Technological utopianism envisions AI as a superior replacement for human judgment. Technological scepticism, conversely, relegates algorithms to mere computational tools, subordinate to human expertise and incapable of capturing the relational dimensions of advice. Both framings prove inadequate for the hybrid reality of contemporary financial organisations. Recent empirical evidence points to a third path: a symbiotic configuration in which AI capabilities and human expertise coalesce into collaborative decision-making systems that, when appropriately designed, can outperform either component operating alone (Dellermann et al., 2021; Bianchi and Brière, 2024). Hybrid advisory arrangements are rapidly emerging as the prevailing configuration in wealth management (Cardillo and Chiappini, 2024). However, this paradigm shift demands a robust regulatory and informational substrate. The proposed European Financial Data Access (FiDA) Regulation (European Commission, 2023) constitutes such a substrate. Building on PSD2, FiDA extends mandatory data sharing to mortgages, savings, investments, pension rights and non-life insurance. It introduces authorisation dashboards, financial data sharing schemes with multi-stakeholder governance, and a new entity, the Financial Information Service Provider (FISP). Drawing on capital structure theory (Modigliani and Miller, 1958, 1963; Myers, 1984; Jensen and Meckling, 1976) and post-crisis banking literature (DeAngelo and Stulz, 2015; Gropp and Heider, 2010), we contend that AI-powered advisory deployed within the FiDA framework can act as a mechanism for rebalancing the funding mix of financial institutions, transforming clients from creditors into stakeholders and thereby enhancing institutional resilience. Two research questions guide our analysis. First, how does FiDA enable a third narrative of human-machine collaboration distinct from automation and

scepticism? Second, how can AI-augmented advisory contribute to a structural rebalancing of bank funding? Each is addressed in turn below.

2 Theoretical Framework and Literature Review

This section reviews four interconnected literatures: AI in financial advisory; human-machine collaboration; FiDA and Open Finance; and capital structure theory in banking. Their intersection grounds our thesis that AI-mediated advisory under FiDA can reshape bank funding mixes.

2.1 Artificial Intelligence in Financial Advisory Services

Artificial intelligence has been described as a prediction technology whose principal economic effect is to drastically reduce the cost of forecasting outcomes from data (Agrawal et al., 2018). In financial advisory, this translates into machine-learning systems capable of inferring investor risk tolerance, allocating assets, and detecting behavioural anomalies at scale (Duan et al., 2019; Judijanto et al., 2022). Recent reviews document a sharp acceleration of AI adoption in financial services since 2020, with robo-advisors emerging among the fastest-growing applications (Cardillo and Chiappini, 2024).

Empirical work has examined whether algorithmic platforms substitute for or complement human advice. Brenner and Meyll (2020) show that robo-advisor use is associated with a 15-16 percentage point reduction in the likelihood of consulting a human advisor, particularly among investors concerned about fraud or conflicts of interest. Aristei and Gallo (2026) refine this picture using Italian household data: robo-advice substitutes for commission-based human advisors but complements independent fee-only advisors, suggesting that the substitution effect is conditional on the incentive structure of the human counterpart. A parallel literature on AI-based recommendation systems shows that hybrid architectures combining content-based and collaborative filtering, often built on gradient boosting (Friedman, 2001), outperform either method alone on banking datasets.

2.2 Human-Machine Collaboration: Narratives, Complementarity, and Trust

Conceptually, the move from automation to augmentation rests on the prediction-judgment dichotomy articulated by Agrawal et al. (2018, 2025): as prediction becomes cheap, the marginal value of complementary inputs (data, judgment, action) rises. Dellermann et al. (2021) extend this insight into the framework of hybrid intelligence, in which humans provide contextual understanding, creativity, and ethical reasoning, while algorithms contribute scalable pattern recognition. Empirical work converges on the finding that hybrid teams can outperform either humans or AI alone, but only when task allocation, interface design and trust calibration are deliberately engineered (Bianchi and Brière, 2024). Three conceptual framings recur. Technological utopianism holds that AI will progressively replace human advisors. Technological scepticism emphasises algorithmic aversion (Dietvorst et al., 2015) and the irreducibility of human judgment. The third, human-machine symbiosis, is empirically supported: D'Acunto et al. (2019) show robo-optimisers improve diversification; Rossi and Utkus (2024) document that robo-advice reduces home bias and fees, lowers cash positions from 19% to 2%, and increases Sharpe ratios; Bianchi and Brière (2024) find that algorithmic alerts raise investor attention while leaving control with the investor, with only marginal additional gains from full automation.

2.3 FiDA, PSD2, and the Open Finance Trajectory

The European Union has progressively moved from sector-specific data access (PSD2 for payment accounts) toward a comprehensive Open Finance framework. PSD2 mandated free access to payment-account data via APIs but produced fragmented standards and weak incentives for data holders to invest in interfaces (European Commission, 2022). The proposed FiDA Regulation (European Commission, 2023) addresses these limitations by extending the scope of mandatory data sharing to mortgages, savings, investments, pension rights and non-life insurance, and by introducing a contractual scheme-based access model with reasonable compensation (Borowicz, 2025).

Three architectural innovations are particularly consequential for AI advisory. First, authorisation dashboards give consumers real-time control over which third parties can access which data and for which purpose (Article 8 FiDA). Second, financial data sharing schemes, industry-developed but governed under equal representation of data holders and users, define technical standards, dispute resolution and compensation models (Articles 9-10). Third, the FISP, authorised under Title V, can access customer data without per-request electronic authorisation, subject to professional indemnity insurance and supervision. From a knowledge-management

perspective, FiDA reframes siloed transactional records as structured, interoperable knowledge assets recombinable across institutional boundaries.

At the time of writing, FiDA has not yet been formally adopted: the European Parliament adopted its first-reading position on 18 April 2024, the Council its general approach on 4 December 2024, and trilogues have been ongoing since April 2025. Political agreement is plausibly achievable during the 2026 Cypriot Presidency, with operational application unlikely before 2028-2030 given the staged timelines envisaged. The European Central Bank (2024) and the European Data Protection Supervisor (2023) have issued opinions emphasising prudential clarity and strict subordination to the General Data Protection Regulation respectively (OECD, 2023).

2.4 Capital Structure Theory in Banking

Capital structure theory provides the analytical lens through which we interpret the macro-financial implications of AI advisory. Modigliani and Miller (1958, 1963) established that, under frictionless markets and absent taxes, firm value is independent of the debt-equity mix; the introduction of corporate tax deductibility of interest then generates a tax advantage of debt. Myers (1984) subsequently contrasted the static trade-off view with the pecking order theory, while Jensen and Meckling (1976) had earlier formalised the role of agency costs.

Banks are structurally distinct. Much of their debt consists of deposits performing liquidity services, and they operate under explicit capital regulation and safety nets. Gropp and Heider (2010) show that capital regulation and mispriced deposit insurance materially shape leverage choices for banks close to minimum requirements. DeAngelo and Stulz (2015) argue that high bank leverage can be socially and privately optimal once the value of producing safe, information-insensitive liquid claims is recognised. Post-crisis regulation requires a minimum 4.5% CET1 ratio, 6% Tier 1, and 2.5% capital conservation buffer, alongside a 3% leverage ratio and the Liquidity Coverage and Net Stable Funding ratios (BCBS, 2021). Alkhazali et al. (2024) find a positive association between pre-crisis regulatory capital ratios and 2020 return on assets, while accounting equity-to-assets shows no such effect, highlighting the role of high-quality capital. The implication is that any mechanism shifting bank funding from short-term deposits toward equity-like instruments tends to enhance resilience without compromising profitability.

3 Conceptual Framework

This section develops the conceptual framework that organises the paper's argument. We proceed in four steps: first, we formalise the three narratives that characterise the discourse on AI in financial advisory; second, we describe the role of FiDA as the informational and trust substrate that operationalises the symbiosis; third, we present an integrated model in which client data, algorithmic insight, and advisor expertise are jointly coordinated as knowledge assets; fourth, we derive two distinct implementation scenarios. Figure 1 summarises the resulting architecture.

3.1 Three Narratives of AI in Financial Advisory

Building on Section 2, we formalise three narratives that organise the discourse on AI in financial advisory. Technological utopianism extrapolates from the prediction-cost reduction of AI to forecast the eventual displacement of human advisors, reading partial substitution effects (Brenner and Meyll, 2020) as the leading edge of full automation. Technological scepticism, conversely, draws on algorithmic aversion (Dietvorst et al., 2015; Burton et al., 2019) and the persistence of relational capital to confine AI to an instrumental role. Both positions share a binary framing of human and machine as substitutes.

The third view, that of human-machine symbiosis, rejects this binary in favour of a complementarity logic grounded in the prediction-judgment distinction (Agrawal et al., 2018, 2025). Where prediction is cheap and accurate, machines should perform it; where judgment is scarce, contextual or value-laden, humans should retain authority. The boundary is neither fixed nor exogenous: task allocation, interface design, trust calibration and incentives jointly determine where complementarity materialises (Dellermann et al., 2021; Bianchi and Brière, 2024). Field evidence on hybrid advisory (D'Acunto et al., 2019; Rossi and Utkus, 2024) is most consistent with this configuration, which we adopt for the remainder of the paper.

3.2 The FiDA Substrate: Data, Trust, and Architecture

Symbiosis requires a substrate. Without high-quality, comprehensive and trustworthy data, machine-learning systems cannot generate insights worth a human advisor's attention; without auditable governance,

advisors cannot legitimately rely on algorithmic outputs in fiduciary contexts. FiDA supplies this substrate along three dimensions: informational, governance, and trust.

The informational dimension produces a cross-domain view of the household balance sheet by extending mandatory data sharing across mortgages, savings, investments, pensions and non-life insurance. The governance dimension is realised through financial data sharing schemes that coordinate technical standards and compensation models under explicit equality between data holders and data users, mitigating the risk that incumbents capture the access infrastructure (Borowicz, 2025).

The trust dimension is operationalised through authorisation dashboards that make consent granular, monitorable and revocable, while FISP authorisation under Title V subordinates data access to professional indemnity insurance and ongoing supervision. Collectively, these provisions translate latent transactional data into structured assets that are intangible, recombinable and governable.

3.3 An Integrated Model of Hybrid Advisory under FiDA

We propose an integrated model, illustrated in Figure 1, in which three knowledge assets are jointly coordinated: (a) client data, made interoperable and consented through FiDA mechanisms; (b) algorithmic insight, generated by machine-learning systems trained on this data and oriented toward suitability assessment, product matching and risk profiling; and (c) advisor expertise, which interprets, contextualises and where appropriate overrides algorithmic outputs in light of client-specific objectives, regulatory requirements and ethical considerations. The coordination of these assets is an ongoing organisational capability requiring governance mechanisms (audit trails, override logs), interface mechanisms (explanation layers, confidence intervals, escalation paths) and learning mechanisms (feedback loops).

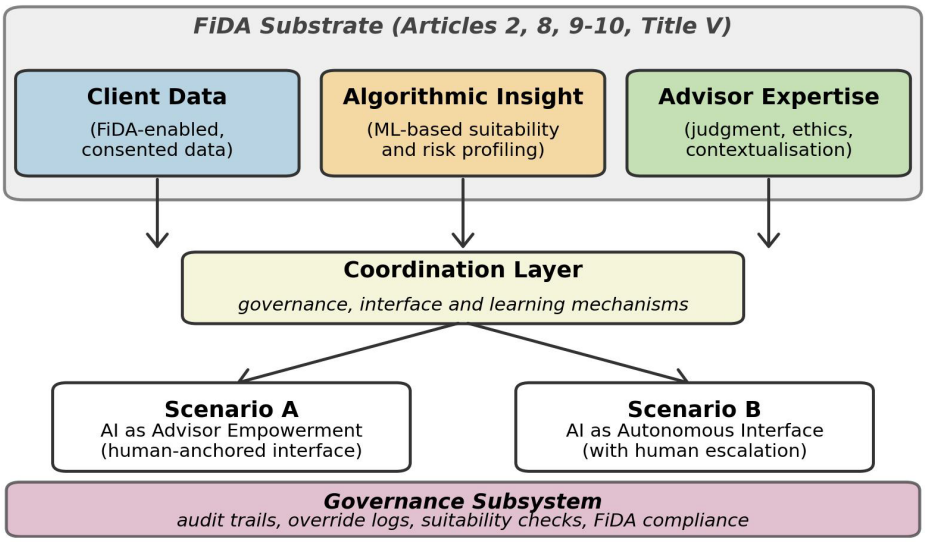


Figure 1: The integrated model of hybrid advisory under FiDA

3.4 Two Implementation Scenarios

The integrated model accommodates two distinct implementation scenarios, summarised in Table 1.

Scenario A (AI as advisor empowerment) has AI generate preliminary product recommendations, suitability scores and risk-flag alerts that are routed to human advisors. The advisor remains the primary point of contact with the client, leveraging algorithmic outputs to allocate time toward relationship building, complex planning and ethical judgment.

Scenario B (AI as autonomous interface with human escalation) has the AI system interact directly with the client for routine onboarding, risk profiling, asset allocation and rebalancing, while predefined triggers (low confidence scores, anomalous behaviour, high-stakes transactions, regulatory thresholds) escalate to human advisors. This configuration maximises scalability and lowers unit costs. Even in Scenario B, humans remain in the loop as governance actors whose oversight is structurally embedded.

Table 1: Two implementation scenarios

Dimension	Scenario A: Empowerment	Scenario B: Autonomous Interface
Primary client interface	Human advisor	AI system
AI role	Decision support	Direct interaction with escalation
Best-fit segments	Affluent, complex situations	Mass-retail, digital-first
Trust profile	Human-anchored	Algorithm-anchored, governed

The scenarios are not mutually exclusive: an institution may operate them in parallel for different segments. What unites them, distinguishing the configuration from both utopian automation and sceptical instrumentalisation, is the deliberate coordination of human and machine contributions through the FiDA-enabled knowledge architecture.

4 Methodology and Proof-of-Concept

This section presents the methodology used to translate the framework into an operational proof-of-concept. The exercise is anticipatory: FiDA is still in the legislative pipeline and authentic cross-domain data are not yet available.

4.1 Research design

We adopt a design-science research approach (Hevner et al., 2004) in which a working artefact is constructed to demonstrate how the integrated model can be operationalised. Evaluation criteria are threefold: alignment with FiDA-mandated data categories and consent mechanisms; generation of suitability-grounded recommendations that augment rather than supplant human judgment; and interpretability and auditability sufficient for advisor oversight.

4.2 Proof-of-concept architecture

The artefact comprises three layers mirroring the three knowledge assets of Section 3. The data layer ingests FiDA-aligned client data spanning the six categories enumerated in Article 2 of the proposed Regulation (credit and account products; savings and investments; occupational and individual pension rights; non-life insurance; firms' credit assessment data); it normalises continuous variables and encodes categorical variables via one-hot transformation. The reasoning layer hosts a supervised machine-learning model that takes client-product pairs as inputs and outputs a binary suitability classification with a confidence score; in an earlier thesis-level proof-of-concept (Rizzo, 2024), a gradient-boosting classifier (Friedman, 2001) was used as a tractable baseline, but the architecture treats the estimator as substitutable. The interface layer presents recommendations along with explanation artefacts and confidence intervals.

Cross-cutting all three layers is a governance subsystem recording every recommendation, override and consent change, providing the audit trail that FiDA's authorisation dashboard and supervisory architecture envisage. This subsystem differentiates the configuration from a black-box robo-advisor and renders the human-machine collaboration organisationally credible.

4.3 Data, variables, and validation approach

Because authentic FiDA data are not yet available, the proof-of-concept is trained on a synthetic dataset designed to mimic the distributional properties and structural categories that FiDA will eventually expose. Two parallel datasets are generated: a client dataset capturing demographic, financial and risk-profile attributes, and a product dataset describing representative financial products across the six FiDA categories. Client-product pairs are labelled for suitability through expert-informed rules grounded in suitability assessments analogous to those required under Article 25 of MiFID II.

The validation protocol follows standard supervised-learning practice and is reported here in methodological terms, since metrics obtained on synthetic data would not generalise to deployment. The dataset is split through stratified sampling; class imbalance is addressed via Synthetic Minority Over-sampling within a pipeline preventing information leakage; hyperparameters are selected by stratified k-fold cross-validation; performance is measured on a held-out test partition using accuracy, ROC-AUC, precision, recall and F1. Beyond predictive metrics, the artefact is assessed against the three design-science criteria of Section 4.1.

4.4 Limitations of the Proof-of-Concept

Three limitations apply. First, synthetic data cannot fully replicate the joint distribution FiDA will eventually expose. Second, expert-informed labelling embeds the labellers' heuristics. Third, the artefact instantiates a single estimator; comparative evaluation across alternative families is left to subsequent work.

5 Macro-Financial Implications

This section develops the macro-financial implications of widespread AI advisory adoption under FiDA, from the bank funding problem to a demand-side rebalancing mechanism, the behavioural conditions for it to materialise, and caveats.

5.1 The funding mix problem in banking

Banks operate at substantially higher leverage than non-financial firms because debt, especially deposits, is the funding instrument best suited to producing safe, information-insensitive liquid claims (DeAngelo and Stulz, 2015). This is privately and socially optimal in normal times, but adverse macroeconomic shocks translate disproportionately into losses for highly leveraged institutions. Post-crisis regulation has responded primarily through supply-side mechanisms (BCBS, 2021); banks with higher pre-crisis capital absorbed COVID-19 losses more effectively (Alkhazali et al., 2024). What remains under-explored is the symmetric demand-side mechanism: how households' allocation choices between deposits and risk capital instruments shape, in aggregate, the composition of bank funding.

5.2 AI Advisory as a demand-side mechanism

Field evidence on robo-advisory documents that AI-mediated advice systematically reduces idle cash positions (from 19% to 2% in one large-scale study, Rossi and Utkus, 2024), increases diversification, and modestly raises equity exposure. When scaled across the FiDA-enabled ecosystem, household wealth currently parked in deposits can be redirected, through suitable recommendations, toward risk capital instruments.

The redirection operates through three channels, summarised in Table 2. First, when AI advisory facilitates subscription of equity, AT1, or subordinated debt issued by the same banking group, deposits convert directly into Tier 1 or Tier 2 capital (Avdjiev et al., 2017; Bormann and Bräuer, 2018). Second, when recommendations steer clients toward third-party risk capital intermediated by the bank, deposits decline and fee income partially substitutes for net-interest margin. Third, even within the deposit envelope, reallocation toward longer-dated instruments improves the maturity profile and Net Stable Funding Ratio.

Table 2: Three channels of demand-side rebalancing

Channel	Mechanism	Effect
Direct conversion	Subscription of bank-issued equity/AT1/subordinated debt	Direct CET1/AT1/Tier 2 increase
Intermediated reallocation	Subscription of third-party risk capital via the bank	Deposit reduction; fee income substitutes margin
Maturity rebalancing	Reallocation toward longer-dated deposit instruments	Improved maturity profile; lower NSFR pressure

5.3 From depositors to stakeholders: behavioural mechanisms

Recommendation alone does not change behaviour; for the macro-financial channels of Section 5.2 to materialise, clients must act on algorithmic suggestions. Choice architecture (Thaler and Sunstein, 2008) is therefore central. Defaults, framing and feedback materially shift adoption: a recent field experiment with a robo-advisor found that switching the default from conventional to sustainable investing raised adoption from 23% to 42%, with 93% of investors retaining their initial setting months later (Hornuf et al., 2025). Earlier work shows that robo-advisory access mitigates the disposition effect, although specific interface choices may attenuate or reinforce it (Back et al., 2023). The same algorithmic recommendation can therefore produce very different aggregate flows depending on its presentation, defaults and rebalancing logic. Holding total assets fixed, an AI advisory system that reallocates a fraction of deposit balances into the bank's own equity-like instruments raises equity capital and lowers deposit funding, reducing the leverage ratio. Following Modigliani and Miller (1958, 1963), in frictionless markets this reallocation would not affect firm value; once frictions, distress costs and tax considerations are introduced (Myers, 1984; Jensen and Meckling, 1976), and once the

convexity of distress costs in adverse states is taken seriously, lower leverage translates into a smaller probability of losses sufficient to threaten solvency.

The transformation of depositors into stakeholders is therefore not merely a metaphor but a measurable change in claim structure. Deposit holders are short-term creditors whose claims must be honoured at par; equity holders absorb downside through reduced distributions and, in the limit, principal write-downs. After a suitability-grounded recommendation accepted under FiDA-compliant consent, a portion of the household's previous deposit balance has been converted into an equity-instrument position whose return co-moves with bank performance. In aggregate, this co-movement realigns private incentives with the institutional resilience that supply-side regulation alone cannot guarantee.

5.4 Resilience, welfare, and caveats

Four caveats temper the argument. First, redirection of deposits into risk capital is welfare-enhancing only insofar as recommendations are genuinely suitable: a system that pushes risk capital onto risk-averse or financially illiterate clients to relieve a bank's funding pressure would breach both fiduciary duty and FiDA's consent and purpose-limitation principles. The architecture of Section 3 explicitly subordinates the recommendation to a suitability test, with human oversight preserved across both implementation scenarios. Second, intra-group recommendations generate a structural conflict of interest requiring governance safeguards, transparency obligations, and reliance on independent FISPs as intermediaries. Third, retail distribution of bank capital instruments raises macro-prudential concerns. The Italian episodes of 2015-2017 (Banca Etruria, Monte dei Paschi di Siena) and the 2023 Credit Suisse AT1 write-down (Banca d'Italia, 2024) illustrate that contingent capital can be wiped out under regulatory action even when held by retail investors. The ESRB (2024) monitors household exposure to bank capital instruments as a vector of systemic vulnerability. Any framework promoting deposit-to-risk-capital transitions must therefore embed product governance, suitability assessments, and concentration limits. Fourth, macro-financial benefits are contingent on aggregation: only when AI advisory is widely deployed and FiDA scheme infrastructure is fully operational do the channels of Section 5.2 produce systemic effects. Subject to these caveats, demand-side rebalancing through AI advisory complements the supply-side capital pressures of Basel III/IV: each addresses one half of the leverage problem, and the joint effect is greater resilience to adverse shocks at lower aggregate cost than either mechanism alone.

6 Discussion

This section addresses organisational, practical, and policy implications, and acknowledges boundary conditions.

6.1 Organisational narratives and professional identities

Adopting symbiosis is not neutral. Where utopianism leads to under-investment in advisor expertise and scepticism to under-investment in algorithmic capability, the symbiotic configuration requires deliberate co-investment in both. Institutions internalising this logic redefine the advisor's role from analyst to interpreter and orchestrator, with implications for hiring, training and incentives (D'Acunto and Rossi, 2023).

6.2 Practical Implications for Financial Institutions

For institutions, the framework suggests a three-track agenda: FiDA readiness (data, consent, audit infrastructure); advisory redesign (Scenario A, B, or segmented); and capital strategy (recognising AI advisory as a potential lever in long-run funding-mix rebalancing while embedding suitability safeguards).

6.3 Policy Implications

Three policy points follow. Supervisors should treat FiDA, MiFID II suitability, and bank capital regulation as a coordinated whole. Macro-prudential authorities should monitor aggregate household exposure to bank-issued risk capital. FiDA authorisation dashboards and FISP supervision should incorporate behavioural-finance considerations (Hornuf et al., 2025).

6.4 Limitations and boundary conditions

The analysis is primarily conceptual; empirical validation requires live FiDA data. The macro-financial argument is stylised and partial-equilibrium; full assessment requires general-equilibrium modelling. The paper focuses on the EU; comparative analysis with the UK and Australia is referenced but not developed.

7 Conclusions

The convergence of AI and FiDA creates the conditions for a third narrative of AI in financial advisory, human-machine symbiosis, with macro-financial consequences that have received insufficient attention. By coordinating client data, algorithmic insight and advisor expertise as knowledge assets, financial institutions can deliver personalised advice while contributing to a structural rebalancing of their funding mix. Through suitability-grounded, FiDA-compliant recommendations, depositors can be progressively transformed into stakeholders, complementing the supply-side resilience pressures of Basel III/IV. Three avenues for future research follow: empirical validation once FiDA enters application; comparative analysis with the UK and Australia; and behavioural research on client responses to AI-mediated narratives, especially under conflict-of-interest disclosures.

8 Ethics Declaration

No ethical clearance was required. The work is conceptual and methodological, involving no human subjects, identifiable personal data, or experimental procedures requiring institutional ethical review. The proof-of-concept data are entirely synthetic.

9 AI Declaration

AI tools were used selectively in line with the IFKAD AI Tools Use Policy: (i) for literature review (Section 2), LLM-based assistants were used to identify candidate sources, subsequently verified by the authors; (ii) for language editing, AI assistants refined sentence structure. No AI-generated content was incorporated without human review and accountability. The conceptual framework, methodology, macro-financial argument and proof-of-concept are the original work of the authors.

References

Agrawal, A., Gans, J. and Goldfarb, A. (2018) Prediction Machines: The Simple Economics of Artificial Intelligence. Boston: Harvard Business Review Press.

Agrawal, A., Gans, J. and Goldfarb, A. (2025) "Machine Intelligence and Human Judgment", Finance and Development, IMF, June.

Alkhazali, O., Helmi, M.H., Mirzaei, A. and Saad, M. (2024) "The impact of capital on bank profitability during the COVID-19 pandemic", Global Finance Journal, 62, 100994.

Aristei, D. and Gallo, M. (2026) "Financial literacy, robo-advising, and the demand for human financial advice: Evidence from Italy", Journal of Behavioral and Experimental Finance, 49, 101125.

Avdjiev, S., Bogdanova, B., Bolton, P., Jiang, W. and Kartasheva, A. (2017) "CoCo issuance and bank fragility", BIS Working Papers, No. 678.

Back, C., Morana, S. and Spann, M. (2023) "When do robo-advisors make us better investors? The impact of robo-advisory on the disposition effect", The European Journal of Finance.

Banca d'Italia (2024) "The Credit Suisse case and AT1 instruments", Questioni di Economia e Finanza, No. 875. Rome: Banca d'Italia.

Basel Committee on Banking Supervision (BCBS) (2021) Basel III: Finalising post-crisis reforms - Holistic evaluation report. Basel: Bank for International Settlements.

Bianchi, M. and Brière, M. (2024) "Human-Robot Interactions in Investment Decisions", Management Science, 72(1), pp. 14-31.

Bormann, J. and Bräuer, S. (2018) "Contingent convertible bonds: Who invests in European CoCos?", Applied Economics Letters, 25(4), pp. 234-238.

Borowicz, M.K. (2025) "Competition, standards, and EU data spaces", working paper.

Brenner, L. and Meyll, T. (2020) "Robo-advisors: A substitute for human financial advice?", Journal of Banking and Finance, 121, 105969.

Burton, J.W., Stein, M.K. and Jensen, T.B. (2019) "A systematic review of algorithm aversion in augmented decision making", Journal of Behavioral Decision Making, 33(2), pp. 220-239.

Cardillo, G. and Chiappini, H. (2024) "Robo-advisors and digital wealth management", Journal of Financial Services Research, 62(1), pp. 91-125.

Centre for European Policy Studies (CEPS) (2016) The continuing saga of Banca Monte dei Paschi di Siena. Brussels: CEPS.

D'Acunto, F., Prabhala, N. and Rossi, A.G. (2019) "The promises and pitfalls of robo-advising", Review of Financial Studies, 32(5), pp. 1983-2020.
D'Acunto, F. and Rossi, A.G. (2023) "Robo-advice: transforming households into rational economic agents", Annual Review of Financial Economics, 15, pp. 543-563.
DeAngelo, H. and Stulz, R.M. (2015) "Liquid-Claim Production, Risk Management, and Bank Capital Structure", Journal of Financial Economics, 116(2), pp. 219-236.
Dellermann, D., Calma, A., Lipusch, N., Weber, T., Weigel, S. and Ebel, P. (2021) "The Future of Human-AI Collaboration", Information Systems Frontiers, 23, pp. 47-65.
Dietvorst, B.J., Simmons, J.P. and Massey, C. (2015) "Algorithm aversion", Journal of Experimental Psychology: General, 144(1), pp. 114-126.
Dietvorst, B.J., Simmons, J.P. and Massey, C. (2018) "Overcoming algorithm aversion", Management Science, 64(3), pp. 1155-1170.
Duan, Y., Edwards, J.S. and Dwivedi, Y.K. (2019) "Artificial intelligence for decision making in the era of Big Data", International Journal of Information Management, 48, pp. 63-71.
European Central Bank (2024) Opinion of the European Central Bank of 30 August 2024 on a proposal for a regulation on a framework for Financial Data Access (CON/2024/29). Frankfurt am Main.
European Commission (2022) Public consultation on the review of PSD2 and on open finance. Brussels.
European Commission (2023) Proposal for a Regulation on a framework for Financial Data Access. COM(2023) 360 final. Brussels.
European Data Protection Supervisor (2023) Opinion 38/2023 on the Proposal for a Regulation on a framework for Financial Data Access. Brussels.
European Systemic Risk Board (ESRB) (2024) ESRB Annual Report 2024. Frankfurt am Main.
Friedman, J.H. (2001) "Greedy function approximation: A gradient boosting machine", Annals of Statistics, 29(5), pp. 1189-1232.
Gropp, R. and Heider, F. (2010) "The Determinants of Bank Capital Structure", Review of Finance, 14(4), pp. 587-622.
Hevner, A.R., March, S.T., Park, J. and Ram, S. (2004) "Design Science in Information Systems Research", MIS Quarterly, 28(1), pp. 75-105.
Hornuf, L., Merkle, C. and Zeisberger, S. (2025) "Nudging investors towards sustainability: A field experiment with a robo-advisor", working paper.
Jensen, M.C. and Meckling, W.H. (1976) "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure", Journal of Financial Economics, 3(4), pp. 305-360.
Judijanto, L., Asfahani, A., Bakri, A.A., Susanto, E. and Kulsum, U. (2022) "AI-Supported Management through Leveraging Artificial Intelligence", Journal of Artificial Intelligence and Development, 1(1), pp. 59-68.
Modigliani, F. and Miller, M.H. (1958) "The Cost of Capital, Corporation Finance and the Theory of Investment", American Economic Review, 48(3), pp. 261-297.
Modigliani, F. and Miller, M.H. (1963) "Corporate Income Taxes and the Cost of Capital: A Correction", American Economic Review, 53(3), pp. 433-443.
Myers, S.C. (1984) "The Capital Structure Puzzle", Journal of Finance, 39(3), pp. 575-592.
OECD (2023) Shifting from open banking to open finance, in OECD Business and Finance Outlook 2023. Paris: OECD Publishing.
Rizzo, G. (2024) Machine Learning e Trattamento dei Dati Finanziari: Prospettive, Regolamentazione e Un Caso d'Uso Esemplificativo. Bachelor's thesis, LUISS Guido Carli University, Rome.
Rossi, A.G. and Utkus, S.P. (2024) "The diversification and welfare effects of robo-advising", Journal of Financial Economics, 157, 103869.
Thaler, R.H. and Sunstein, C.R. (2008) Nudge: Improving Decisions about Health, Wealth, and Happiness. New Haven: Yale University Press.

Public Policies and Knowledge Democracy in Innovation Ecosystems: A Quintuple Helix Analytical Framework

Leani Lauermann Koch[1] , Aline Luiza Brusco Pletsch[1] , Graciele Tonial[1] , Florinda Matos[2], Beatriz Carla Koch[3]

[1] Universidade do Oeste de Santa Catarina - UNOESC, Chapecó-SC, Brazil

[2] (DINÂMIA'CET), Instituto Universitário de Lisboa - Iscte, Lisboa, Portugal

[3] University for Continuing Education Krems, Krems an der Donau, Áustria

Abstract

This article proposes an analytical framework of the Quintuple Helix for Knowledge Democracy-driven Innovation Ecosystems through Public Policies, as a theoretical-conceptual structure for analysing how innovation ecosystems may transition from functional efficiency towards socio-ecological legitimacy, inclusion, and resilience. The instrument is grounded in the integration of five theoretical fields: innovation ecosystems, helix-based innovation models, systemic public innovation policies, knowledge democracy, and sustainability-oriented governance. The framework is structured around five interdependent helices: Government, Academia, Industry, Civil Society, and the Environment. These helices are articulated through systemic public policies, understood as mechanisms of coordination, mediation, integration, and institutional orchestration. Its defining characteristic is the central position of Knowledge Democracy, conceived as the symbolic and operational core where knowledge flows, social legitimacy, participation, institutional learning, decision-making authority, and the collective appropriation of knowledge converge. The framework is operationalised through definitions, supporting theoretical references, and guiding questions organised by helix, intended to direct empirical investigations across different territorial contexts. The principal contribution of the article is to reposition innovation ecosystems as complex territorial arrangements, shaped by asymmetries, coordination processes, political choices, and knowledge governance. From a policy perspective, the instrument suggests that innovation strategies must combine directionality, participation, inter-helix articulation, socio-environmental criteria, and institutional feedback. As a theoretical-conceptual contribution, the article offers an analytical structure comprising propositions for future empirical applications across different territorial contexts.

Keywords: Quintuple Helix Innovation Systems, Knowledge Democracy, Systemic Innovation Policies, Innovation Ecosystems Governance, Socio-ecological Transition

Paper type: Academic Research Paper

1 Introduction

Innovation ecosystems help to address this complexity by framing innovation as a relational, non-linear, and territorially embedded process, involving heterogeneous actors — including government, academia, industry, civil society, and other institutional arrangements. However, when ecosystems are analysed primarily through the lens of market coordination, entrepreneurial density, or competitiveness, their democratic and socio-ecological dimensions frequently remain underdeveloped.

Helix-based innovation models offer an important pathway for overcoming this limitation. The Triple Helix accounts for innovation through the interaction between university, industry, and government (Etzkowitz and Leydesdorff, 2000). The Quadruple Helix expands this structure by incorporating civil society, culture, media, and users, connecting innovation to the knowledge society and to knowledge democracy (Carayannis and Campbell, 2009, 2014). The Quintuple Helix extends this logic further by adding the natural environment as a constitutive dimension of innovation systems, linking knowledge production to sustainability and socio-ecological transition (Carayannis and Campbell, 2010; Carayannis, Barth and Campbell, 2012).

This theoretical evolution alters the analytical significance of innovation. Within the Quintuple Helix, innovation is understood as a socially constructed, territorially embedded, and ecologically conditioned process. Knowledge Democracy is central to this discussion, as it emphasises epistemic pluralism, equitable access to knowledge, social participation, and the collective appropriation of knowledge in decision-making processes (Campbell, 2019). Accordingly, innovation ecosystems must be evaluated not solely on the basis of their technological or economic outcomes, but also in terms of their capacity to integrate diverse forms of knowledge, reduce asymmetries, and support sustainable development.

Public policies play a structuring role in this process. Traditional innovation policy frequently treats the state as a corrective actor responding to market failures, whereas more recent approaches conceive of public policy as a directional and market-shaping force, capable of mobilising resources, coordinating actors, and orienting

innovation towards collective ends (Mazzucato, 2016). Public investment in innovation may also produce persistent economic effects, stimulating research and development and broadening productive capacities (Ciaffi, Deleidi and Mazzucato, 2024). Public policies must therefore be understood as mechanisms of orchestration, coordination, and institutional learning within innovation ecosystems.

Notwithstanding these advances, the literature still lacks integrated conceptual models that connect public policies, Quintuple Helix interactions, Knowledge Democracy, and socio-ecological sustainability. This article is guided by the following research question: *How can systemic public policies structure Quintuple Helix innovation ecosystems so as to strengthen Knowledge Democracy, inclusion, and socio-ecological sustainability?* In response to this question, the article proposes and discusses an analytical framework of the Quintuple Helix for Knowledge Democracy-driven innovation ecosystems. The instrument is structured around government, academia, industry, civil society, and the environment, encompassing social legitimacy, participation, decision-making authority, and institutional learning. The article contributes to the field by shifting the analysis of innovation ecosystems from efficiency and interaction towards legitimacy, inclusion, sustainability, and institutional learning.

2 Theoretical Foundations

2.1 Innovation ecosystems and public policies

Innovation ecosystems are increasingly understood as complex, relational, and territorially rooted arrangements. They involve actors, activities, institutions, artefacts, and relationships that influence the innovative performance of a given actor or group of actors (Granstrand and Holgersson, 2020). This perspective broadens the analysis of innovation beyond isolated firms and linear technology transfer, emphasising interaction, interdependence, and collective learning.

Recent literature reinforces this systemic interpretation. Carayannis et al. (2018) argue that ecosystems may be interpreted through helix architectures, given that innovation depends upon dynamic, multi-sectoral, and multi-level interactions. Gifford et al. (2021) likewise demonstrate that knowledge-intensive innovation ecosystems evolve through the co-evolution of entrepreneurial activity and innovation policies. This indicates that ecosystems are shaped not only by bottom-up initiatives, but also by public coordination, institutional experimentation, and policy directionality.

Public policies are fundamental to this process. Traditional innovation policy frequently frames the state as a corrective actor that intervenes when markets fail. However, this view is insufficient to account for how innovation trajectories are shaped within complex ecosystems. Mazzucato (2016) argues that public policies may evolve from a focus on market correction towards a focus on market creation, actively shaping technological directions, mobilising resources, and generating new opportunities for public value. Accordingly, innovation policy is not merely an external support mechanism; it is a structuring force.

Systemic public policies coordinate actors, reduce asymmetries, develop institutional capacities, and support learning throughout the ecosystem. They may also connect innovation to broader social challenges, such as sustainability, inclusion, and territorial cohesion. Ciaffi, Deleidi and Mazzucato (2024) provide evidence that public investment in research and development can produce persistent macroeconomic effects and stimulate private research and development. This reinforces the notion that public action can influence innovation capacity beyond short-term financing.

From this standpoint, innovation ecosystems require more than entrepreneurial density or technological infrastructure. They demand governance mechanisms capable of aligning knowledge production, productive capacities, social demands, and environmental priorities. The Organisation for Economic Co-operation and Development (OECD) also emphasises that innovation policies are increasingly dependent on coordination, directionality, and systemic governance to address complex social challenges (OECD, 2023). For this reason, public policies must be understood as mechanisms of orchestration, mediation, and institutional learning within innovation ecosystems.

2.2 From the Triple Helix to the Quintuple Helix

Helix-based models provide an important theoretical foundation for analysing the institutional architecture of innovation. The Triple Helix model accounts for innovation through the interaction between university, industry, and government (Etzkowitz and Leydesdorff, 2000). In this model, universities contribute to knowledge production, firms transform knowledge into products and services, and governments furnish regulatory,

institutional, and political support. The Triple Helix was therefore instrumental in advancing the understanding of innovation in knowledge-based economies.

Nevertheless, the Triple Helix presents limitations when analysing innovation in relation to social legitimacy, participation, and democratic governance. Its focus on university–industry–government relations may under-represent users, communities, culture, media, and broader social demands. The Quadruple Helix addresses this limitation by incorporating civil society, media and cultural publics, users, and democratic contexts into innovation systems (Carayannis and Campbell, 2009; Carayannis and Campbell, 2014).

The Quadruple Helix alters the significance of participation in innovation. Civil society is no longer merely a passive recipient of technological outcomes. It becomes a source of knowledge, legitimacy, demand articulation, and social orientation. In this respect, innovation is linked to the knowledge society and to knowledge democracy, insofar as epistemic pluralism and the diversity of knowledge become conditions for more legitimate innovation systems (Carayannis and Campbell, 2014).

The Quintuple Helix extends this structure further by incorporating the natural environment as a fifth helix. Carayannis and Campbell (2010) and Carayannis, Barth and Campbell (2012) argue that the natural environment should not be treated solely as an external constraint or as a context affected by innovation. It must also be understood as a driver of knowledge production, innovation, and socio-ecological transition. The Quintuple Helix thus connects innovation to sustainability, ecological responsibility, and long-term social transformation.

This evolution from the Triple Helix to the Quintuple Helix does not represent a mere accumulation of actors. It constitutes a conceptual shift. The analytical focus moves from institutional interaction for economic innovation towards a broader architecture that integrates the knowledge society, democratic participation, and socio-ecological sustainability. Carayannis and Campbell (2021) reinforce this point by arguing that Quadruple and Quintuple Helix innovation systems are grounded in democracy and ecology. Carayannis and Campbell (2022) likewise present the Quintuple Helix as an integrative device for modern, sustainable, and democratic knowledge economies and societies.

This article adopts the Quintuple Helix as its central theoretical architecture. The Quadruple Helix is relevant as an intermediate stage, insofar as it introduces civil society and democratic participation. The analytical focus, however, is the Quintuple Helix, given that the proposed instrument integrates government, academia, industry, civil society, and the environment — a scope essential for analysing innovation ecosystems as sociotechnical and socio-ecological systems.

2.3 Knowledge, democracy, and innovation governance

Knowledge Democracy deepens the democratic dimension of innovation ecosystems. It extends democracy beyond formal political institutions, emphasising epistemic pluralism, equitable access to knowledge, social participation, co-production, and the collective appropriation of knowledge. Campbell (2019) connects democracy, knowledge, and innovation by arguing that democratic quality can enable innovation. From this perspective, innovation is shaped not only by technological capacity, but also by the conditions under which knowledge circulates and is employed in collective decision-making.

This perspective challenges technocratic visions of innovation governance. Innovation cannot be regarded as democratic merely because diverse actors are formally present within the ecosystem. Formal participation differs from substantive participation. Formal participation may entail consultation, presence, or representation without real influence. Substantive participation requires that actors be in a position to shape agendas, priorities, resource allocation, and policy learning.

Bunders et al. (2010) argue that Knowledge Democracy requires conditions under which both dominant and non-dominant actors contribute meaningfully to public policies and research agendas. This distinction is particularly significant in innovation ecosystems. Universities, firms, and governments frequently concentrate technical authority, institutional influence, and financial resources. In the absence of democratic governance mechanisms, civil society may be included symbolically, whilst actual decision-making remains concentrated within technocratic or market-oriented spheres. Seeley et al. (2019) reinforce this argument by highlighting the importance of dialogue, the recognition of diverse voices, and the challenging of dominant knowledge hierarchies.

Democratic governance of innovation is therefore distinct from technocratic coordination. Technocratic coordination may enhance efficiency, but does not necessarily broaden legitimacy, inclusion, or collective learning. Democratic governance requires institutional arrangements that enable plural forms of knowledge to influence decisions. It also requires accountability mechanisms, participatory spaces, and feedback loops through which innovation outcomes may inform future public policies.

The distinction between symbolic inclusion and real influence is equally important. Civil society and environmental concerns may appear in policy discourse without effectively shaping priorities. Within the Quintuple Helix, however, civil society and the environment must not be peripheral or merely rhetorical. They are structural dimensions of innovation. Their incorporation is substantive only when it genuinely influences the direction, legitimacy, and sustainability of innovation trajectories.

Knowledge Democracy is therefore the analytical core that connects knowledge flows, social legitimacy, participation, institutional learning, and decision-making authority within innovation ecosystems. It is the principle that transforms the Quintuple Helix from a multi-sectoral structure into a democratic governance framework.

3 Conceptual and Methodological Approach

This article is theoretical-conceptual in nature. It presents neither empirical results, nor interview data, nor comparative findings. Its objective is to construct and discuss an analytical framework that integrates theoretical fields generally addressed in isolation, offering an operational instrument for empirical investigations into the governance of innovation ecosystems in the light of Knowledge Democracy. The conceptual construction of the instrument draws upon an integrative review of the literature on innovation ecosystems, helix-based models, public innovation policies, knowledge democracy, sustainability, and innovation governance. These fields provide the theoretical foundation for understanding innovation ecosystems as complex sociotechnical and socio-ecological arrangements.

The first theoretical field concerns innovation ecosystems. This literature sustains the understanding of innovation as a relational and territorially embedded process, shaped by actors, institutions, artefacts, and knowledge flows (Granstrand and Holgersson, 2020). It also enables the article to approach ecosystems as governance spaces, rather than merely as environments for entrepreneurship or technological performance.

The second field concerns helix-based innovation models. This article traces the transition from the Triple Helix to the Quadruple and Quintuple Helix frameworks, with emphasis on the Quintuple Helix as the central architecture (Etzkowitz and Leydesdorff, 2000; Carayannis and Campbell, 2010; Carayannis, Barth and Campbell, 2012). This theoretical field enables the framework to articulate government, academia, industry, civil society, and the environment.

The third field concerns public innovation policies. This article adopts the perspective that public policies are not confined to correcting market failures, but may shape markets, define trajectories, coordinate actors, and orient innovation towards public ends (Mazzucato, 2016). This perspective grounds the role of systemic public policies as structuring mechanisms within the framework.

The fourth field concerns Knowledge Democracy and the democratic governance of innovation. This literature enables the instrument to incorporate epistemic pluralism, equitable access to knowledge, social participation, co-production, and the collective appropriation of knowledge (Campbell, 2019; Bunders et al., 2010). It also provides the basis for distinguishing formal participation from substantive participation, and knowledge diffusion from knowledge co-production.

The fifth field concerns sustainability and socio-ecological transition. The Quintuple Helix literature frames the natural environment as a structural component of innovation systems, rather than as an externality (Carayannis, Barth and Campbell, 2012). This supports the inclusion of the natural environment both as a helix and as a condition for sustainable innovation trajectories.

The methodological contribution of this article thus resides in the conceptual integration of these five fields and in the construction of a structured analytical framework. The article does not seek to conduct a systematic review in the technical sense, nor does it present empirical procedures such as interviews, coding, or comparative analysis. The empirical application of the instrument may be developed in future studies, across different territorial and institutional contexts.

4 Quintuple Helix Analytical Framework for Innovation Ecosystems

Building upon the articulation of the five theoretical fields presented above, this section proposes an analytical framework designed to guide empirical investigations into the governance of innovation ecosystems in the light of Knowledge Democracy. The framework does not constitute a normative or prescriptive model. It is an analytical instrument that enables the systematisation, operationalisation, and comparison of how different territories articulate the five helices, systemic public policies, and the democratic processes of knowledge production and appropriation.

The presence of five helices does not presuppose that such presence automatically produces democratic innovation. The quality of interaction amongst the helices depends upon whether knowledge is accessed

equitably, co-produced, and collectively appropriated. This condition is precisely what the framework's guiding questions seek to render analytically visible. In this respect, Knowledge Democracy is not an additional concept incorporated into the Quintuple Helix structure, but rather the principle that transforms a multi-sectoral architecture into a democratic governance framework.

Three complementary dimensions organise the instrument, as set out in Figure 1. The first identifies each helix, namely Government, Academia, Industry, Civil Society, and the Natural Environment, and presents its definition within the context of innovation ecosystems. These definitions delimit the analytical role of each helix and establish the criteria by which its functioning may be investigated empirically. The second presents the reference authors underpinning the role of each helix in the international literature, conferring upon the framework both theoretical anchorage and traceability for future applications. The third dimension addresses guiding questions grouped into four thematic blocks: government policies for innovation; the role of universities and society; civil society participation; and aspects of public policies and research and development investment.

The thematic blocks were deliberately organised to capture the interdependence amongst the helices. Innovation ecosystems frequently exhibit fragmentation and weak alignment amongst actors, and the mere presence of universities, firms, government agencies, and civil society organisations does not guarantee systemic integration. Coordination requires public agendas, institutional capacity, and spaces for collective learning (Edquist, 1997; Lundvall, 2010). For this reason, Block 1 investigates government policies as the institutional conditions that structure interaction amongst the remaining helices. Block 2 examines how universities operate in knowledge production, mediation, and transfer in partnership with other actors. Block 3 analyses civil society participation, distinguishing between formal presence and the capacity to influence agendas and priorities. Block 4 investigates research and development investments as mechanisms capable of generating systemic effects when combined with institutional coordination and socio-environmental criteria (Ciaffi, Deleidi and Mazzucato, 2024). This interdependence amongst the blocks reflects the relational logic of the Quintuple Helix and precludes isolated, sector-specific readings of individual helices.

The inclusion of the Natural Environment as a constitutive helix follows from the understanding that the environment is not an external variable or a passive context affected by innovation. It is a dimension of innovation systems that conditions technological trajectories and orients knowledge production towards sustainability and socio-ecological transition (Carayannis and Campbell, 2010; Carayannis, Barth and Campbell, 2012). Block 4 therefore enables investigation of the extent to which innovation policies incorporate socio-environmental criteria, and the degree to which ecosystem outcomes are evaluated beyond economic performance.

Positioned at the centre of the framework, Knowledge Democracy signals that the assessment of innovation ecosystems cannot be restricted to indicators of technological or economic performance. It must also capture the distribution of epistemic power: who defines research problems, who accesses knowledge resources, who influences public policy priorities, and who benefits from innovation outcomes. The role of the state, in this context, is not limited to financing. It involves building institutional capacities, creating inter-helix coordination mechanisms, supporting participatory spaces, and ensuring that innovation outcomes feed back into institutional learning (Mazzucato, 2016). The guiding questions were constructed to render these dimensions analytically operational, enabling the identification, within the ecosystems under investigation, of both inclusion mechanisms and the structural asymmetries that constrain substantive participation.

Helices	Definitions of the Helices	References
Government	The "Government Helix" represents the institutional and government definiition of the Stata within the Innovation Helix models. This boha highlights the oneffierxent use of the government is a regolator. Siavever, uan its duroer of his sonere and eccegmete, ecting as one of the righrires and afflore to drive innovation and sustainable development. Through policixs, regulations, and programs, by government procenss collaboration between adademia, industry, and society, creating a rooduzive environment for innovation and tecknological advancement. Its rele as a cslalyst fents to create condtfams for aztive ndleenal program in a restiprehesseas and collaborative socress, focevttating protective, educating popressics, and cetllisating strategies that beees innovation and economic development commatity.	Cal & Latta, 2022; Carayansia & Canaphell, 2099;2019,2021; Carayansia & Morewaka, 2022, Carayansia et al., 2012, 2017, 2020, 2021a, 2021b, 2021a; 20112, 2022, Forwa Eccnien et al., 2022, Forum Econômica Mundial (2017); Malik et al., 2017, Reasan & Fellalveka, 2021, Sebitte et al., 2019, Shin et al., 2022.
Academia	The "Adafomia Helix" appreends the dynamic and collaborative interaction between higher education institutions, society, and the productive sector, engineerring the generation of knowledge, chrer development, and technology martiks. In this centent, sucvensclon ant as knowledge production eanters, pervesting ecneration through imonitrergieraty reresrk and strategic perrecallogs with innovation, governnaere and wril sacress. The collaboration drives social, econersic, and socie-technical development. Entering sustainability and ceaspstifivenssa, creating a fiverable environment for the cevreetion of innovative solutions and the advanvement of a knowledge-based and demostatic society.	Amkili & Faliohadar, 2022, Cai & Latta, 2022, Arakki et al., 2012, Maiik et al., 2022, Darito Foroum et al., 2022, Mali et al., 2022, Foruin Econômica Mondial, 2022, Corayaasia & Marmaska, 2022, Carayanala & Caraphall, 2019, 2019, 2022, Carsynasia et al. 2012, 2012, 2012, 20118, 20113, 20113, 20214, 2022.
Industry	The "Industry Helis" represents for cyeergy between the private vetate and nanssweers, playing a baselopnasal rale as test-dnreang centurate knowledge inta innovative dynamic and secwers. This distantion msads not for in collaborative and directive approach. focused on the practical application of knowledge to promote ecomaina martsh, premots compahictessus, and drive asstainable growth. The interaction between the ecadncrue helie and the industrial lallix constvrera to creative, technology trenties, and collaboration innovative forms, conmitituting in the development of more agile and relattive innovation ecosystems.	Amkil et al., 2010, Cai & Latta, 2022, Carayanala & Carapmala, 2017, Carayanala & Caraphall, 2099, 2010, 2017, Curaynasia et al., 2012, 2017, 2013, 20314, 2021a, 2021a, 2021a, 2022, Dusta-Reusean et al., 2022, Forom Econômica Mundial, 2017, Malik et al., 2017, Baruan & Fellihabes, 2012, Shin et al., 2022.
Civil Context	The "Civil Society Helix" is a dynamic and collaborative structure that inereporates the active and davere partizipation of cifinoes, non-governmental organizations, and commundity, createng out for its empleses on increroas, social responsibility, and honore-og-engagement. This kelis is distsegarized in the dolety to presvent the incrrertten of innovative seletions, and more shiltlc policies and innovation practices, and drive sustainable development through the interaction between different reccerd acress, as well as the public and private serters.	Amkil et al., 2010, Cai & Lattu, 2022, Carayanala & Caraphall, 2099, 2017, Carayanala et al., 2013, 2017,2013, 2011a, 2011e, 20214, 20214, 2022, Carayanaia & Marunaka, 2023, Forum Romeas et al., 2022, Forum Econômico Mundial, 2013, e 91, Malik et al., 2017, Remore & Felihhake, 2017, Schiltz et al., 2019, Shin et al., 2022.
Natural Environment	The "Natural Environment Helix" is the innovation model integrates environmental exstensity and espaste into economic development, prescnvting environmental enstaieability. Sent the devaception in the implementation of solutions in the market. This belistia appreach is engresse the eust expendence between innovative, knowledge, and the natural environment, emphorizing the need to preserve natural resources and adget materiality practices to proceste environments, a naxural economy, and climate change nesigatisa, incosperating the natural environment in an adfferenet hega neccess and impreasions are developed responsisty, balancing technological progress with environmental cobeseration. Collaboration is repreted to create sustainable solutions that respect the planet's ecological lanrts and promote a scose balanced and hrmacutous future.	Amkil et al., 2010, Carayansia & Caraphell, 2019, 2011, Carayansia & Morawaka, 2017, Carayansia et al., 2017, 2017, 2011a, 2011a, 2021a, 2022, Forum Romare et al., 2022, Forum Económico Mundial, 2022.

Questions

Block 1 - Government Public Policies for Innovation

This block explores the role of public policies in encreraging innovation, challenges and orgentanities of implementation, and thatt condestation in the innovation esysteters.

- Do you know of any nxisting public policies that encosrage innovation in your regiox? What are the most relevant strangtirs?
- How do you evaluate the government coles in artiftiating policy policies aneed of innovation?
- How can yas cuaanat regulation gnoductly condicues to aducating nancratans in the neropeana?
- Do police you can neccatige the participouts of different incest in innovative activities teest in workshops, hacksftkans, and other arreas? Who are the mast participasts?
- Is your gracesm, in the technical capacity of public policy realaze sefflarves in effiericy processs innovation?
- What are the loggest challenges lailed is the mplementation of tublic renovvierx systems?
- Are there differences in smplementing public innovation policies as cartant used in the economy, education, or social arres? Can you provide strangle?
- How can government satise at different levels (level, regional, and ar nealty consatlons in arongiltening innovation exeyrement?

Block 2 - Role of Universities/Sociotechnical

This block addresses the fundessesaal role of universites in developing one valums and technologies, as well as how they colllictions with the productive etates.

- How have universities ocntributed to the development of vrzy graduate and technological processes that velataes eceouscies' evmperticensent?
- How do you evaluate asovemnent oate to strium intection and knowledge provessend in pertsanlniy with the product os assur?
- How do you are have government provessnsus ds bstared them eafflouass io hamd comsyocent?
- What are the wara challenges loved by conpacere in balancing technoscian and human asernted innovaction in the Quiztsple Helis Innovation ceuponer?
- How can universities play a more active rele in developing innovement and technologies that prossses sestainable development?
- How can proeness beass as avnlasea and selivery netrevas to pevetert ponctinl rostals in ocaas of processses? What could de improved?
- How do you are anccessary role in uvecesiesing technological anereetions into more applicable market acvocrer?
- What opprenesates can mécero has explaes to strengtoce thra relationship with the prodactive acsess within the context of the Quixtaple Helie?

Block 3 - Social Context

This block exploses how civil coniety one countions in the innovation ecesystem, strengthening collaboration between different and shalders and helping to solve social challenges.

- How can civil soniety activity participate in the Quizteple Helix innovation ecssystem by producing solations to social challenges such as digital relations and sigalt aurers in technology?
- What are the impasct of civil activity participation in innovation and in improving the quality of products and services offered by companize?
- What are massa have su lect innovation poticies cougnus iy nilamned social and pcasities mist challenge? Can you pece entergare?
- What are the nuax benefite that innovative policies provats in maaaly and the market in prorcss?
- How can collabert line between civil ercaede, ecademis, and government generete more reclinesa and socially responsible uxcent hear?
- Hev we had faet sead erremy has eeld and coffcones to shaging innovation policies that impacst acconsity irie?
- How can the participanse of NDOs and sesial eccccuneat between the impact of public innovation policies in reslernamed arces?
- Do you know how can the private ezrret moce effectively engags seall ustility in developing innovations that meel ar axicle?

Block 4 - Socioeconomic Context

This block deals with the impacts of public policies and investement in R&D on economic development and explame the socioeconomic challenges related to innovation.

- How can public policies and R&D inverse rate drive sustainable economic growth is to plans with different socioeconomic callities?
- Have the respesees allevated to research and development noovative programs, with collaborative projects between tudceses and industry, generally significant ecosurch, regnent?
- What a portogame, have the recosory and the lack of financial resources have strateries in the technological adveroement of companies? If es, in wher most?
- What is the importance of collaboration between government, academic, industry, society, and the environment in the Condogia Hulir madel in in ies usachenen?
- What can be moot challenges leved in cnccciating the different castes for the effective implementation of public innovative policiss?
- How has the government roontal invovlems in ssamficat occecsttus and ocesanric development? Are these exemples of eccures?
- Is your spantan, have public policies been different in procesting maininability and social responsibility io the current socioeconomic context?
- How have innovation policies increased mangerrtile competecemens and devecatied the economy? Can you partfly with enemyles?

Figure 1 — Quintuple Helix Analytical Framework for Knowledge Democracy-driven Innovation Ecosystems.
Source: Authors' own elaboration.

6 Conclusion

The analytical framework presented in this article operationalises, through definitions, theoretical references, and guiding questions organised by helix, a structure for investigating how innovation ecosystems articulate democratic governance, public policies, and socio-ecological sustainability.

Within this framework, public policies are not interpreted as external interventions triggered by market failures. They are mechanisms that define agendas, coordinate actors, align instruments, reduce asymmetries, and create institutional conditions for collective learning. The role of the state is not to supplant the remaining helices, but to orchestrate the conditions for inter-helix cooperation, support participatory spaces, and ensure that innovation outcomes feed back into institutional learning.

Positioned at the centre of the framework, Knowledge Democracy furnishes the criterion for analysing the democratic quality of innovation governance. It renders visible the distinction between formal and substantive participation, between knowledge diffusion and knowledge co-production, and between symbolic inclusion and real influence over decision-making. Innovation may be technically advanced and economically successful whilst remaining socially exclusionary or territorially unequal. Treating knowledge as a collective resource — accessible, co-produced, and appropriated by a plurality of actors — is the condition that distinguishes ecosystems with democratic governance from those with governance that is only formally plural.

The instrument contributes to the literature on innovation ecosystems by interpreting them as complex territorial arrangements shaped by actors, institutions, power asymmetries, knowledge flows, and political choices. It contributes to Quintuple Helix theory by reinforcing civil society and the environment as structural dimensions of innovation governance. It also contributes to the public policy literature by demonstrating that innovation policies are not solely corrective or financial instruments, but may equally structure trajectories, build institutional capacities, and orient innovation towards public value.

For policymakers and practitioners, the framework indicates that innovation policy must combine strategic directionality with democratic governance. Public action should promote inter-helix coordination, support participatory spaces, reduce territorial and social asymmetries, incorporate socio-environmental criteria, and create feedback mechanisms for institutional learning. In this sense, innovation policy becomes not merely a tool for competitiveness, but also a mechanism for democratic and sustainable development.

The article has limitations. Its contribution is theoretical and conceptual, and it presents no empirical testing of the instrument. Future studies may apply the framework empirically across different territorial contexts,

comparing how innovation ecosystems vary in terms of public governance, inter-helix articulation, Knowledge Democracy, sustainability, and systemic feedback. Such applications may help to refine the instrument and evaluate its utility for analysing innovation ecosystems across diverse institutional, social, and socio-ecological contexts.

7 Acknowledgements

The authors gratefully acknowledge the financial support of the Santa Catarina Research and Innovation Foundation (*Fundação de Amparo à Pesquisa e Inovação do Estado de Santa Catarina* — FAPESC) and the Coordination for the Improvement of Higher Education Personnel (*Coordenação de Aperfeiçoamento de Pessoal de Nível Superior* — CAPES). This support contributed to the development of the broader research programme from which this article derives.

8 Ethics Declaration

This article is theoretical-conceptual in nature and draws exclusively upon the review and integration of academic literature. Accordingly, ethical approval was not required for the development of this study.

9 AI Declaration

The authors made use of ChatGPT, developed by OpenAI, to assist with language refinement and text condensation. The AI tool was not employed to produce empirical data, generate research findings, or substitute theoretical analysis. All outputs produced with AI assistance were reviewed, edited, and validated by the authors, who assume full responsibility for the content, arguments, references, and final version of the article.

References

Bunders, J.F.G. *et al.* (2010) "How can transdisciplinary research contribute to knowledge democracy?," in *Knowledge Democracy*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 125–152.

Campbell, D.F.J. (2019) *Global quality of democracy as innovation enabler: Measuring democracy for success*. 1st ed. Cham, Switzerland: Springer International Publishing.

Carayannis, E.G. *et al.* (2018) "The ecosystem as helix: an exploratory theory-building study of regional co-opetitive entrepreneurial ecosystems as Quadruple/Quintuple Helix Innovation Models: The ecosystem as helix," *R and D Management*, 48(1), pp. 148–162. Available at: https://doi.org/10.1111/radm.12300.

Carayannis, E.G., Barth, T.D. and Campbell, D.F.J. (2012) "The Quintuple Helix innovation model: global warming as a challenge and driver for innovation," *Journal of Innovation and Entrepreneurship*, 1(1), p. 2. Available at: https://doi.org/10.1186/2192-5372-1-2.

Carayannis, E.G. and Campbell, D.F.J. (2009) "'Mode 3' and 'Quadruple Helix': toward a 21st century fractal innovation ecosystem," *Journal international de la gestion technologique [International Journal of Technology Management]*, 46(3/4), p. 201. Available at: https://doi.org/10.1504/ijtm.2009.023374.

Carayannis, E.G. and Campbell, D.F.J. (2010) "Triple Helix, Quadruple Helix and Quintuple Helix and how do knowledge, innovation and the environment relate to each other?: A proposed framework for a trans-disciplinary analysis of sustainable development and social ecology," *International Journal of Social Ecology and Sustainable Development*, 1(1), pp. 41–69. Available at: https://doi.org/10.4018/jsesd.2010010105.

Carayannis, E.G. and Campbell, D.F.J. (2014) "Developed democracies versus emerging autocracies: arts, democracy, and innovation in Quadruple Helix innovation systems," *Journal of Innovation and Entrepreneurship*, 3(1). Available at: https://doi.org/10.1186/s13731-014-0012-2.

Carayannis, E.G. and Campbell, D.F.J. (2021) "Democracy of climate and climate for democracy: The evolution of Quadruple and Quintuple Helix innovation systems," *Journal of the Knowledge Economy*, 12(4), pp. 2050–2082. Available at: https://doi.org/10.1007/s13132-021-00778-x.

Carayannis, E.G. and Campbell, D.F.J. (2022) "Towards an Emerging Unified Theory of helix architectures (EUTOHA): Focus on the Quintuple Innovation Helix Framework as the integrative device," *Triple Helix (Heidelberg)*, pp. 1–11. Available at: https://doi.org/10.1163/21971927-bja10028.

Ciaffi, G., Deleidi, M. and Mazzucato, M. (2024) "Measuring the macroeconomic responses to public investment in innovation: evidence from OECD countries," *Industrial and Corporate Change*, 33(2), pp. 363–382. Available at: https://doi.org/10.1093/icc/dtae005.

Edquist, C. (1997) "Systems of innovation approaches: their emergence and characteristics"," in C. Edquist (ed.) *Systems of Innovation: Technologies, Institutions and Organizations*. London: Pinter, pp. 1–35.

Etzkowitz, H. and Leydesdorff, L. (2000) "The dynamics of innovation: from National Systems and 'Mode 2' to a Triple Helix of university–industry–government relations," *Research Policy*, 29(2), pp. 109–123. Available at: https://doi.org/10.1016/s0048-7333(99)00055-4.

Gifford, E., McKelvey, M. and Saemundsson, R. (2021) "The evolution of knowledge-intensive innovation ecosystems: co-evolving entrepreneurial activity and innovation policy in the West Swedish maritime system," *Industry and Innovation*, 28(5), pp. 651–676. Available at: https://doi.org/10.1080/13662716.2020.1856047.

Granstrand, O. and Holgersson, M. (2020) "Innovation ecosystems: A conceptual review and a new definition," *Technovation*, 90–91(102098), p. 102098. Available at: https://doi.org/10.1016/j.technovation.2019.102098.

Lundvall, B.-Å. (2010) "User-producer relationships, national systems of innovation and internationalisation," in B.-A. Lundvall (ed.) *National Systems of Innovation*. London: Anthem Press, pp. 47–70.

Mazzucato, M. (2016) "From market fixing to market-creating: a new framework for innovation policy," *Industry and Innovation*, 23(2), pp. 140–156. Available at: https://doi.org/10.1080/13662716.2016.1146124.

Mazzucato, M. (2018) "Mission-oriented innovation policies: challenges and opportunities," *Industrial and Corporate Change*, 27(5), pp. 803–815. Available at: https://doi.org/10.1093/icc/dty034.

OECDE (2023) *Reviews of Innovation Policy: Korea 2023*. Edited by OECD Reviews of Innovation Policy. Paris: OECD Publishing.

Seeley, J. *et al.* (2019) "Creating a space for global dialogue on knowledge democracy: experiences from the inaugural global assembly for knowledge democracy," *Educational Action Research*, 27(1), pp. 22–39. Available at: https://doi.org/10.1080/09650792.2018.1552170.

Wang, Q.-J. *et al.* (2021) "The impacts of democracy on innovation: Revisited evidence," *Technovation*, 108(102333), p. 102333. Available at: https://doi.org/10.1016/j.technovation.2021.102333.

The Organizational Challenges of Digital Transformation at SMEs

Anikó Almási
Obuda University - John von Neumann Faculty, Budapest, Hungary

Abstract

Research on SME competitiveness shows that it is not the level or strength of individual competitiveness factors that determines the success of SMEs, but rather the capabilities—particularly managerial skills—that empower or limit the implementation of these factors within the organization. Digitalization became essential factor of success of enterprises. While digital transformation has become a critical driver of competitiveness, its outcomes remain uneven, particularly among SMEs, where the limited resources, fragmented information flows and resistance to change constrain the value creation. Contemporary organizations increasingly operate at the intersection of knowledge, human behaviour and digital technologies, where value creation depends not on isolated technological investments but coordinated integration of organizational capabilities, cultural norms and knowledge practices.

This study analyses the determinants of digital transformation (DT) success and failure at small and medium-sized enterprises (SMEs) through an integrated perspective that combines technological, organizational, behavioural and environmental dimensions. Building on established theoretical frameworks, including the Technology-Organization-Environment (TOE) model and GLOBE cultural dimensions model the study develops an extended analytical approach that incorporates cultural and leadership factors. The results reveal a significant gap between technological potential and organizational readiness. Although SME managers increasingly recognize the importance of digitalization and innovation as key competitiveness factors, current digital practices remain strongly control-oriented and fragmented.

The knowledge generated by digital systems often remains centralized, underutilized, or limited to small groups of managers, which restricts its contribution to organizational learning, innovation, and employee empowerment. This behaviour reflects broader cultural patterns characterized by risk aversion, hierarchical decision-making, and a preference for monitoring and cost control over experimentation and knowledge sharing. This paper argues that DT should not be understood as a purely technological process, but as an organization-wide learning and integration phenomenon in which knowledge creation, sharing and utilization are embedded in everyday practices. It is necessary to pay attention to the specific characteristics of SMEs and local cultural aspects.

Keywords: Digital Transformation, SME, success factor, organization

Paper type: Practical Paper

1 Introduction

While digital transformation (DT) becomes a strategic imperative across industries, its outcomes remain uneven – particularly in small and medium-sized enterprises (SMEs). The SME sector generally has limited resources, and this fragmented knowledge flow and behavioural resistance often constrain the realization of business value (Zang et al., 2022, Kahveci, 2025). However, in most countries, SMEs make a significant contribution to GDP and employment. Innovation and digitalization are usually a basic factor of SME competitiveness models (Chikán and Czakó, 2009, (Szerb, 2007), as well as the limitations of leadership knowledge and the management of organizational capabilities.

Organizational and environmental factors have become more significant in DT as well. Mostly the environment is analysed from a narrow perspective, focusing on direct relationships, and look at the impact of the industry, competitors, and customers. This approach often manifests at the organizational level as well, where the focus is limited to digital readiness, change management, and the corporate lifecycle. However, new research suggest that a broader scope is required to examine both organizational and external, especially the cultural factors. The analysis of internal factors is company-specific, but the external environment differs significantly depending on the country. Therefore, to ensure the validity and reliability of the results, it is appropriate to analyse cultural influences on a country-specific, homogeneous basis. This study therefore focuses on the Hungarian SMEs and aims to contribute to the literature by revealing the impact of the external environment and culture on the success of DT.

For this purpose, the research focuses on SMEs in Hungary, as external influences can be adequately analysed in the context of homogeneous external factors.

- Besides technological factors, what other elements contribute to successful DT?
- How does this relate to the specific characteristics of SMEs' resources and competencies?
- What cultural factors are characteristic of Hungary that may influence the DT process and its success?

During the research, the representative database of the “Organization of the Future” research project was used for analysis (Marciniak et al., 2022). Cultural dimensions were analysed based on the GLOBE framework (House et al., 2004). For the extended theory of DT, we used the TOE model (Baker, 2012).

2 Theoretical Background

Digital Transformation (DT) affects the entire company; it cannot be reduced simply to the implementation of a new, data-driven system. On the one hand, DT impacts the external dimension: it changes customer value propositions, contributes to the development of new products and services, and opens new marketing channels for the company (Loonam et al., 2018). The external dimension also serves as a framework: it is the social environment in which the company operates, which often culturally determines potential solutions and obstacles. On the other hand, DT also plays a strong internal role: it can lead to a more innovative company if it is integrated at the organizational level and affects the entire internal dimension of value creation at the operations and process levels (Vogelsang et al., 2018). A completely new mindset at all organizational level, new business processes, updated management knowledge, and leadership roles are essential for the success of DT (Gupta et al., 2020).

At the same time, these same SME characteristics can also be an advantage during digital transformation: processes are less complex, the flow of information and decision-making can be faster and more flexible, and managers have direct knowledge of the resource and competency gaps affecting the organization (Kim and Jin, 2024). Digital transformation often involves reviewing process milestones, optimizing information flow, restructuring the organization, etc. Therefore, digitization also plays a role in professionalization, especially among SMEs. A carefully considered digitization project often results in higher-level corporate operations, thereby raising the company to a higher level of competitiveness (Murphy, 2002).

2.1 Digital Transformation as an organizational reconfiguration

The complexity of DT has now reached a level where we are no longer discussing how tasks are allocated between IT and business, but rather how to digitize the entire organization. This requires strong connectivity and a new management approach on the part of both IT and management. Digitalization does not mean the same thing as it did in the early 2000s, so it is worth revisiting models that analyse the success of basic and widespread IT projects, such as Information Systems Success (ISS) (DeLone and McLean, 1992, Alter, 2008), Critical Success Chains (CSC), Analytic Hierarchy Process (AHP) and Fuzzy Cognitive Maps (FCM) (Rodriguez-Repiso et al., 2007), as their decision support function is underestimated. Many authors apply industry-based models and concentrate on manufacturing companies. In their opinion, manufacturing companies are particularly affected by Industry 4.0 due to the application of technology and are therefore at a different stage of digital maturity than non-manufacturing companies (König et al., 2019, Nwaiwu, 2020). business processes or whether IT serves business model development. Others emphasize the advantage of knowledge-intensive sectors (Gemmell, 2017), as connectivity and learning processes between organizations are more intense, so innovative practices spread more quickly, and the absorption and rapid application of sectoral knowledge can mean survival for SMEs in the initial phase (Krishna, 2018). Benchmarking also typically takes place within the industry, which also increases the role of sectoral customs and practices, accumulated experience, and networks (Garwin et al., 2008).

Instead of listing technological factors, the focus has shifted to complex models that can account for both environmental and organizational factors, and but also including the organizational fit of the technology in the assessment and addressing the (lack of) competence. Baker (2012) summarized and further developed Tornatzky and Fleischer's (1990) TOE framework.

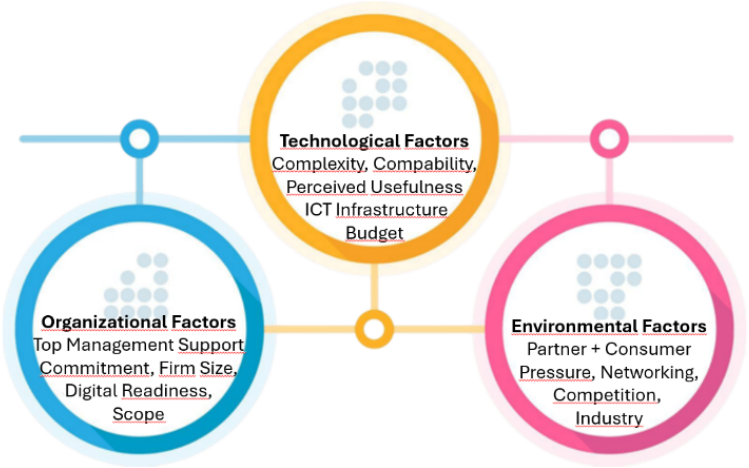


Figure 1: Types of DT / IT adoption factors (based on Baker, 2012)

While companies try to respond the constant external pressure to digitize, the environment is also subject the same trend: existing business models are becoming obsolete, and disruption demands rapid strategic responses from companies (Wirtz et al, 2022). Finding the right answers requires organizational structural changes, which are not covered by the DT project; rather, it requires leadership commitment and the expertise to manage change (Hess et al., 2016). The strength of TOE framework lies in capturing structural factors, but it underestimates micro-level dynamics as cognition, trust and leadership. This study emphasizes that behavioural and cultural dimensions can be mediating variables of explanation of DT, because leaders and culture shape how the hard elements – resources, capabilities – interact.

2.2 Cultural context

Hungary's cultural profile, as identified by the GLOBE study (House et al., 2004), provides important context to understand organizational behaviour in DT process.

Hungary has a relative high power distance: authority is typically concentrated at the top management level, and employees have limited involvement is strategic decisions. Hierarchically structured companies are typically centralized, generally the top managers have the right to make the final decisions, and the authority of middle managers has been reduced.

Strong uncertainty is also remarkable in Hungary. Risk-averse managerial practices, centralized organization structures and top-down information flow is common practice. As opportunities and possibilities increase, so does the role of internal control: managers strive to make processes and resources measurable and controllable in real time (Power, 2007).

The in-group collectivism has significant role in the business environment. Loyalty plays a key role both within the organization and in relationships between companies. Trust and the importance of relationships often outweigh the objective performance indicators. All these specialities tend to encourage group resistance to support or acceptance of change.

3 Methodology

Considering the general and exploratory nature of the research, a mixed-methods approach was adopted, combining secondary research findings with primary research data.

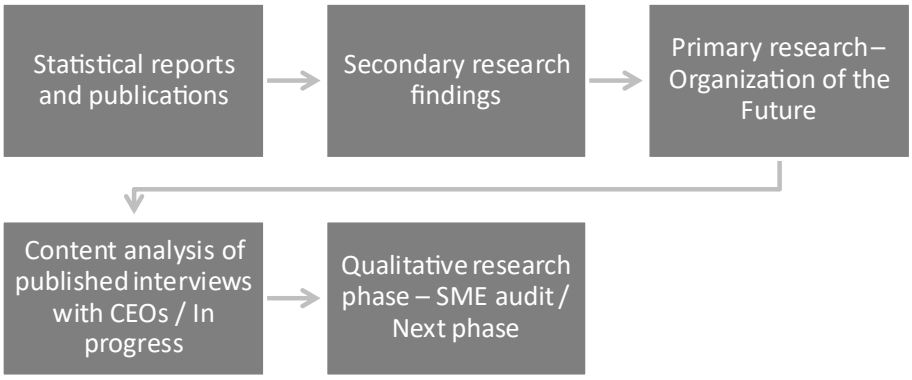


Figure 2: Data collection and analysis design

Organization of the Future research project (Marciniak et al., 2022) involved 305 companies; the survey was conducted by "Scale Research" market research company in March and April 2022. The target group was companies operating in Hungary that are larger than micro and small enterprises, employ 50 or more people. The respondents were senior managers of the companies. 53,5% of the companies are smaller medium sized with 50-99 employees, 29,37% are medium sized and 16,8% of the companies in the sample are large company with 250+ employees. The final database was constructed using stratified random sampling. The strata were defined based on company size, industry, and regional quotas.

4 Findings

4.1 Secondary data on Hungarian SMEs' digital transformation

4.1.1 Statistical reports

Based on the "Administrative burden on enterprises" research and publication of the Hungarian Central Statistical Office (HCSO) Hungarian companies spent 40 205 hours with economic statistics of divisions and 127 818 hours with Statistics of economic organizations in 2024 (HCSO, 2024). Despite this result two-thirds of Hungarian SMEs will have a website by 2023, but this figure drops to only 64% for small businesses with fewer than 50 employees. The content of the websites is also telling most of them contain product descriptions, but only 28.9% use their corporate website for sales (HCSO, 2020). A more popular solution among Hungarian SMEs is the use of social media platforms: 40.5% of businesses have both social media and a website. However, the proportion of digital services is negligible: only 16% of companies offer online customization of products and services, and only 9% had a chatbot, virtual customer service, or online communication options. The level of customer database exploitation is also low: only 13% of businesses track returning customers and tailor their online content to their needs. 93% of SMEs' online activity and 78% of their social media use was product related. Only 47% use online platforms to analyse customer feedback, 27% to collaborate with business partners, and only 12% to recruit new employees. The analysis clearly shows that the online activity of Hungarian businesses is less intense but fundamentally similar in structure to the EU average. The main difference is that SMEs in Hungary are of a single type, while in other EU countries they are more diverse and diversified.

Beyond online sales, however, there are many other areas for digitalization, and Hungary is far behind the EU average in terms of integrating digital technologies. 6.0% of local SMEs performed big data analysis, and 1.0% used external services for this purpose. The corresponding figures for the EU were 13% and 3.0%. Hungarian companies typically analyse geolocation data, and the use of data from their own systems and devices (2.6%) and data generated using social media (2.7%) is also more widespread. This means that Hungarian SMEs are lagging other countries in the region.

4.1.2 Secondary research findings

In their research, Szedmák and Szabó (2025) used semi-structured interviews to survey Hungarian experts with several years of management and digital transformation experience. Among the technological factors, the ten respondents highlighted the unreliable operation and slowness of the systems. They see the reasons for this as being that the digitized system was introduced too late, expectations within the organization were high, but often, due to budget constraints, they were unable to develop, purchase, and implement a system that would meet current needs.

At the operational level, the problem of technology and organizational alignment came to the surface. It often became apparent during the digitization project that the company did not have a deep enough understanding of its own processes. Hierarchical organizational structures encourage divisional egoism, while a digital transformation project requires the entire organization to work together. Ad hoc IT investments (often driven by currently available grant funds) are common, and the systems purchased in this way are also difficult to integrate into the company's operations.

Communication gaps and deficiencies in project management skills also proved to be critical factors in the success of digital transformation. Unprepared and unmeasurable projects are significantly more likely to end in failure, and research has also shown that negative experiences discourage digital projects in the long term. Management competencies include clearly defining the goals of the digital transformation project, as this is also the basis for measuring success.

According to the Digital Entrepreneurship Report / DEE index (Ács et al., 2025) in the case of Hungary, the period under analysis (2017–2022) showed a decline rather than stable convergence. Hungary has strong digital infrastructure, and the country also benefits from strong progress in Digital User Citizenship (DUC) factor. The society is relatively strong in digital participation, user trust, online service adoption, so Hungarian SMEs face less demand-side resistance from customers regarding digital adoption. But this individual level strength does not manifest at the institutional level: despite the strong technology foundation the entrepreneurial conversion is weak. Hungarian SMEs struggle with weak innovation and digital matchmaking and financing, ecosystem development, institutional stagnation, lack of trusted regulatory systems, lower governance coordination and inefficiencies. The authors suggest that it is essential to move from “building infrastructure” to “building entrepreneurial outcome” policy, because the high technology context cannot compensate for weak organizational and environmental conversion.

Another report of the DEE index research (Szerb and Cziegler, 2024) emphasizes the external and cultural background of these findings. The specialities of the Hungarian business environment are the patronage entrepreneurship: state-dependent entrepreneurship, the political favouritism, rent-seeking business behaviour and weakened market competition. Digital and technological infrastructure improved faster than institutions, and related knowledge has not yet reached a level that allows the technology to be effectively used and applied. That is the reason of the gap between technological readiness and entrepreneurial outcomes.

A study (Kárpáti et al., 2024) examined the digital attitudes of Hungarian and Slovak companies based on the DESI index at 210 enterprises. The most significant differences were found in the integration of ERP systems and internal processes, as well as in the goals of digitization: while Slovak companies want to improve the adaptability of their business models with digital solutions, the main motivation for Hungarian companies is cost reduction and resource allocation optimization.

4.2 Research on Organization of the Future

The structured interviews were conducted by interviewers. During the interviews, statements were presented, and respondents were asked to rate their level of agreement with the statements on a 5-point Likert scale. The resulting dataset was analysed using SPSS software. The relationship between the statements and the research questions was evaluated using ANOVA and correlation analysis. The analysis highlights the role of environmental and institutional factors in shaping managerial perceptions and digital behaviour.

The research examined the use of strategy data support, enterprise resource planning (ERP), customer relationship management (CRM), business intelligence systems and reports (BI), automation, and robotization during digitalization. The responding companies believed that the company of the future will be successful and competitive if it is technologically advanced, modern, constantly evolving, and innovative. This was followed by characteristics such as adapting to environmental expectations, being digitally mature, and being capable of renewal. At the end of the list were environmental awareness, timelessness, and a skilled workforce.

Hungarian company executives rated environmental impacts as significantly more influential (4.05 on a 5-point Likert scale) than international studies dealing with the environment. This uncertainty factor poses the greatest challenge for domestic businesses.

Companies that place a strong emphasis on the use of digital technologies reported that their leaders pay attention to errors, failures, and deviations in direct proportion to their level of digital maturity. In contrast, however, companies that consistently lead the way in strategic initiatives have focused less on errors. This also demonstrates that it is not the level of technology, but rather the attitude of leadership that determines how DT is used. When examining managers’ attitudes toward mistakes in relation to the company’s strategic initiative, the following results emerged:

Table 1: Strategic innovation compared with attitude of failure

ANOVA						
		Sum of Squares	df	Mean Square	F	Sig.
Leaders keep track of all mistakes.	Between Groups	15,076	4	3,769	2,599	,036
	Within Groups	414,763	286	1,450		
	Total	429,839	290			
Leaders direct others attention toward failures to meet standards.	Between Groups	9,617	4	2,404	3,300	,011
	Within Groups	214,175	294	,728		
	Total	223,792	298			
Leaders delay responding to urgent questions.	Between Groups	22,426	4	5,607	2,891	,023
	Within Groups	562,422	290	1,939		
	Total	584,848	294			

When examining managers' attitudes toward mistakes in relation to the company's DT level, the following results emerged:

Table 2: Priority to digital technologies compared with attitude of failure

ANOVA						
		Sum of Squares	df	Mean Square	F	Sig.
Leaders keep track of all mistakes.	Between Groups	8,454	4	2,114	1,438	,221
	Within Groups	426,281	290	1,470		
	Total	434,735	294			
Leaders direct others attention toward failures to meet standards.	Between Groups	23,346	4	5,837	8,485	,000
	Within Groups	206,371	300	,688		
	Total	229,717	304			
Leaders delay responding to urgent questions.	Between Groups	26,819	4	6,705	3,437	,009
	Within Groups	575,430	295	1,951		
	Total	602,248	299			

This highlights another important finding of the research: control is vital for Hungarian managers, and often the motivation for digitalization is to be able to monitor their company's operations more conveniently and accurately. Therefore, digitalization is often used not for forward-looking, growth-oriented functions, but to control data and processes. This, in contrast, also highlights the problem of the return on investment of DT: the control function does not increase business value. Access to data is still severely restricted for other members of the organization, and the flow of information remains closed even after digitization.

Hungarian SMEs focus primarily on the digitization of sales channels, often with under-planned, ad hoc digitization projects. In addition to company size, implementation is also a question of money: the immediate application of new technologies is the privilege of the companies with the highest turnover. However, while larger companies can afford to rapidly introduce innovative technologies, smaller companies are more experimental, one reason for which may be their limited resources. The results also show sectoral effects: both the use of new technologies (36%) and experimentation (40%) were most prevalent in the manufacturing industry, with other sectors lagging far behind.

The data analysis revealed that most managers use corporate data assets to evaluate the current situation and identify problems. Data analysis is even less common for the purpose of communicating decisions and monitoring implementation, and this seems to be the privilege of larger SMEs. The low level of integration is also revealed: companies use automation mostly for isolated activities, and only 13% use it comprehensively, with a process-oriented approach.

Data-driven corporate operations can provide opportunities for both centralization and decentralization of decision-making. Quantitative analyses are typically (72.5%) only available to a narrower management layer, and 23.6% of managers completely reject the possibility that data analyses made available through digitalization will become accessible to a wider audience in the coming years. Only 9.2% of companies plan to expand data access within their organization and decentralize their decision-making processes in the next three years. The sample clearly showed the impact of leadership style on the centralization of decision-making processes, but this topic goes beyond the scope of the present study.

5 Discussion

The findings confirms that the success of DT in Hungarian SMEs is determined less by the availability of technology and more by the interaction between organizational capabilities, leadership behaviour and institutional environment. The secondary data show that Hungarian SMEs possess relatively strong basic digital infrastructure. However, the low level of ERP integration, big data analytics reveals that digital maturity remains shallow and functionally isolated. Also the Organization of the future research phase shows many digital projects fail because firms do not completely understand their own processes before attempting digitalization.

The empirical results also highlight the dominance of centralized managerial behaviour. This control-oriented managerial behaviour significantly reduces the knowledge-creation potential of DT. Managers are not necessarily passive but rather controlling reactive. It means that they do not develop proactively and focus on errors, deviations and risk. DT thus serves monitoring and risk reduction rather than decentralization or organization growth.

These findings are strongly connected to the Hungarian cultural context described by the GLOBE dimensions. High power distance supports centralised decision-making and restricted participation, while strong uncertainty avoidance encourages defensive strategies, preference for proven technologies, and low tolerance for experimentation. In-group collectivism strengthens loyalty-based organizational processes and resistance to externally initiated change. Environmental factors further intensify these challenges. This reinforces short-term thinking and risk-averse investment behaviour. DT projects fail not because technology is unavailable, but because organizations lack shared understanding, strategic alignment and cultural readiness for knowledge sharing. This study therefor suggest that successful DT requires a shift from technology-oriented implementation toward organization-wide knowledge integration.

6 Conclusion

The study demonstrates that DT in SME sector cannot be understood as a purely technological project. Because of the rapidly changing innovation and digitalization solutions nowadays the IT knowledge is not enough to manage the corporate digitalization projects. The complexity of contemporary DT requires an organization-wide approach: strong connectivity between the top management and other departments of the companies. Therefore, the success and failure factors of DT have also changed. Alignment between business and functional level IT strategy, organization support, resistant to change and resilience became the important bottleneck of DT. Culture and environmental conditions are closely interconnected to DT. Organizational integration, strategic coordination and the ability to transform digital information into shared organizational knowledge and renewed processes.

However, the cultural characteristics helps to explain why DT often serves managerial supervision. Overestimation of environmental uncertainty and the persistence of centralized managerial attitudes create significant barriers to organizational integration and long-term digital value creation.

For successful DT project Hungarian SME leaders must move beyond isolated IT projects and adopt systematic organizational perspective. This includes stronger process orientation, broader access to data, leadership development, employee involvement and redesign of decision-making practice. DT should be treated not as an IT investment but as a strategic capability-building process.

The practical implication of this study is that future competitiveness depends not only on the quantity of digital tools, but on how effectively firms integrate technology into organizational routines, collaborative processes and strategic renewal. Without decentralization, organization restructure, and managerial support for change even advanced digital technologies will remain underutilized and unable to generate sustainable business value.

7 Acknowledgements

The empirical research “Organization of the future” was supported by Corvinus University of Budapest, Institute of Strategy and Management.

8 Ethics declaration

The empirical research “Organization of the future” was organized by the Institute of Management Science at Corvinus University. Research Director: Róbert Marciniak, PhD. As a member of the research team, the author of this study was responsible for the sections of the questionnaire and the research report covering strategy and competitiveness-inhibiting factors. The rest of the research is the intellectual property of the research group members. Ethical approval was granted by the Research Ethics Committee of Corvinus University:

https://www.uni-corvinus.hu/fooldal/egyetemunkrol/szenatus-egyetemi-testuletek/kutatasetikai-bizottsag/kiadott-kutatasetikai-engedelyek-hatarozatok/#accordion-item-1002

9 AI declaration

AI tool was used in developing the paper: deepl.com for checking translation.

References

Ács, Z., Szerb, L., & Apostol, S. (2025). Digital Entrepreneurship Ecosystem Index Report: The Danube Region. Vienna: Vienna Institute of Global Studies (VIGS).

Alter, S. (2008). Defining information systems as work systems: implications for the IS field. *European Journal of Information Systems*, *17*(5), 448–469.)

Baker, J. (2012): The Technology – Organization – Environment Framework. *Information Systems Theory*. pp. 231-245

Bass, B. M., & Avolio, B. J. (1994). *Improving organizational effectiveness through transformational leadership*. Thousand Oaks, CA: Sage.

Chikán, A. – Czakó, E. (2009): In competition with the world. The competitiveness of our businesses on the threshold of the new millennium. Akadémiai Kiadó, Budapest

DeLone, W. H. and E. R. McLean, "Information systems success: The quest for the dependent variable," *Information systems research*, vol. 3, pp. 60-95, 1992.)

Garvin, D. A.– Edmondson, A. C.– Gino, F. (2008): Is Yours a Learning Organization? *Harvard Business Review* 2008 (03)

Gemmell, R. M. (2017): Learning Styles of Entrepreneurs in Knowledge-Intensive Industries. *International Journal of Entrepreneurial Behavior & Research* 23(3): 446-464.

Gupta, S., Modgil, S., Gunasekaran, A., & Bag, S. (2020). Dynamic capabilities and institutional theories for industry 4.0 and digital supply chain. In 21(3). *Supply chain forum: An international journal* (pp. 139–157). Taylor & Francis

Hess, T., Matt, C., Benlian, A., & Wiesb¨ock, F. (2016). Options for formulating a digital transformation strategy. *MIS Quarterly Executive*, 15(2), 123–139.

House, R. J., Hanges, P. J., Javidan, M., Dorfman, P. W., & Gupta, V. (Eds.). (2004). *Culture, leadership, and organizations: The GLOBE study of 62 societies*. Thousand Oaks, CA. Sage.

Hungarian Central Statistical Office (2024): Administrative burden on enterprises. https://www.ksh.hu/burdens

Hungarian Central Statistical Office (2020): https://www.ksh.hu/kkv/pdf/Infokommunikacios_eszkozok_hasznalata.pdf https://www.ksh.hu/docs/hun/xftp/idoszaki/ikt/2020/02/#gyerekcipbenjradigitlistechnolgikintegrltsga

Kárpáti-Daróczi Judit, Saáry Réka, Tick Andrea: Assessing enterprises' attitudes towards „digitainability" – A comparison of Hungarian and Slovakian characteristics. *Statistical Review*102 : 3 pp. 261-283. , 23 p. (2024)

Kahveci, E. (2025). Digital Transformation in SMEs: Enablers, Interconnections, and a Framework for Sustainable Competitive Advantage. *Administrative Sciences*, 15(3), 107

Kim, J., & Jin, W. (2024). Impact of digital capabilities on entrepreneurial performance in SMEs. *Journal of Innovation & Knowledge*, 9(4), 10060

König, M.– Ungerer, C.– Baltes, G.– Terzidis, O. (2019): Different Patterns in the Evolution of Digital and Non-digital Ventures' Business Models. *Technological Forecasting & Social Change* 146: 844–852.,)

Krishna, H. S. (2018): Entrepreneurial Learning and Indian Tech Startup Survival: An Empirical Investi gation. *Asian Journal of Innovation & Policy* 7(1): 55–78.

Loonam, J., Eaves, S., Kumar, V., & Parry, G. (2018). Towards digital transformation: Lessons learned from traditional organizations. *Strategic Change*, 27(2), 101–109. doi:10.1002/jsc.2185

Marciniak, R. (ed.) (2022): The Organisation and Management of the Future. Research Report. Corvinus University of Budapest. DOI: 10.14267/978-963-503-929-6

Murphy, T. (2002): *Achieving Business Value from Technology*. John Wiley & Sons

Nwaiwu, Fortune; Duduci, Meri; Chromjaková, Felicita; Otekhile, Cathy-Austin: Industry 4.0 concepts within the Czech SME manufacturing sector: An empirical assessment of critical success factors. *Business: Theory and Practice*. 2020, vol. 21, issue 1, p. 58-70 ISSN:1648-0627

Power, M. (2007): *Organized Uncertainty: Designing a World of Risk Management*. Oxford University Press. Oxford

Rodriguez-Repiso, Luis , Rossitza Setchi, Jose L. Salmeron, Modelling IT projects success: Emerging methodologies reviewed, *Technovation*, Volume 27, Issue 10, 2007, Pages 582-594.

Saáry, R., Kárpáti-Daróczi, J., & Tick, A. (2022). Profit or less waste?: Digitainability in SMEs: A comparison of hungarian and slovakian SMEs. *Serbian Journal of Management*, 17(1), 33–49.

Szedmák, B. – Szabó, Zs.R. (2025): Success factors of digital solution implementation. *Budapest Management Review*, 56 (11). pp. 55-70.

Szerb, László and Czigler, Enikő ed. (2024). The digital entrepreneurship ecosystem in the Central Eastern European countries. In: Foresight and STI governance 18 (4), S. 18 – 32.

Szerb L. (2007): Growth opportunities for small and medium-sized enterprises: external and internal supporting and inhibiting factors. *Market Economy Foundation*, Budapest

Tornatzky, L.G. – Fleischer, M. (1990): *The processes of technological innovation*. Lexington Books.

Vogelsang, K., Liere-Netheler, K., Packmohr, S., & Hoppe, U. (2018). Success factors for fostering a digital transformation in manufacturing companies. *Journal of Enterprise Transformation*, 8(1–2), 121–142. doi:10.1080/19488289.2019.1578839

Wirtz, B. W., Weyerer, J. C., & Heckeroth, J. K. (2022). Digital disruption and digital transformation: A strategic integrative framework. *International Journal of Innovation Management*, 26(03), 2240008.

Zhang, X.; Xu, Y.; Ma, L. Research on Successful Factors and Influencing Mechanism of the Digital Transformation in SMEs. *Sustainability* 2022, 14, 2549

Microfoundations of Dynamic Capabilities: Intelligent Knowledge for Sustainable Internationalisation in Higher Education

Judit Fortvingler
Eotvos Lorand University, Budapest, Hungary

Abstract

As global competition intensifies, international higher education institutions must develop strong organisational agility to remain competitive. However, the sector's traditional emphasis on volume-driven massification frequently generates bureaucratic rigidities that constrain adaptation within international operations. While dynamic capabilities theory addresses such environmental volatility, its predominantly macro-level focus offers limited insight into how universities adapt internally.

To address this gap, this paper examines the cognitive microfoundations of dynamic capabilities, within the context of internationalisation, focusing on how the intelligent knowledge of international academic leaders facilitates strategic renewal and sustainable internationalisation. Building on this perspective, the paper proposes a conceptual framework that links individual-level cognition to organisational-level adaptation.

The research adopts a qualitative, comparative methodology, examining three structurally distinct Hungarian universities governed by different operational models: a traditional state university, a centralised public institution, and a market-oriented private university. Data were collected via in-depth interviews with senior leaders responsible for international affairs. Using the proposed framework, the analysis maps leaders' cognitive capabilities onto the key stages of dynamic capabilities: sensing, seizing, and transforming.

The findings suggest that managerial cognition plays a critical role in enabling agility in international higher education. Process-oriented expertise allows executives to detect emerging global shifts, while experience-based judgement helps reconcile financial pressures with academic values under conditions of uncertainty. In addition, contextual awareness allows leaders to interpret external demands in ways that reduce internal resistance during organisational change. The results also indicate that governance structures significantly shape these processes: while market-oriented institutions rely on rapid structural adjustments, state-funded universities depend more heavily on individual leadership to overcome institutional inertia and build organisational resilience.

From a practical perspective, the study highlights the importance of developing multidimensional cognitive competencies among international academic leaders, rather than relying solely on formal administrative procedures. To ensure continuity in internationalisation strategies, individual expertise should be embedded within institutional processes and practices.

By introducing a micro-level perspective on dynamic capabilities, the paper shows how managerial cognition can support sustainable internationalisation in institutions outside the global elite tier. The study contributes to a more context-sensitive understanding of academic agility, demonstrating that institutional resilience in international operations can be built without uncritically importing corporate management models.

Keywords: dynamic capabilities, microfoundations, intelligent knowledge, higher education internationalisation, organisational resilience.

Paper type: Academic Research Paper

1 Introduction

Internationalisation has evolved from a peripheral activity into a central strategic priority, driven by global complexity and the need for sustainable growth (de Wit, 2020; Knight, 2004). As knowledge-intensive organisations, universities must adapt to shifting demographics, funding, and regulations (Seeber et al., 2016). Ensuring strategic sustainability requires more than traditional administrative management: it demands profound organisational agility and adaptive resilience.

While dynamic capabilities (DCs) theory explains organisational adaptation (Teece, Pisano and Shuen, 1997; Teece, 2007), its application to higher education is problematic. Universities face a unique tension: the systemic drive to build volume through massification (Marginson, 2016) creates bureaucratic rigidities that conflict with the agility required for DCs. Since macro-level models fail to explain how this tension is navigated daily, a shift toward the microfoundations of DCs (Helfat and Peteraf, 2015) is required, focusing on the role of individual academic leadership in driving internationalisation.

To understand how this tension is resolved at the micro-level, this study introduces intelligent knowledge as a theoretical bridge, building on Schiuma's (2011) knowledge assets framework and Davenport and Prusak's (1998) domains of human expertise. By examining the cognitive and practical tools of boundary-spanning

international leaders, the paper addresses three research questions: (RQ1) How do international academic leaders apply intelligent knowledge within institutional sensing, seizing, and transforming processes? (RQ2) How does institutional context shape these dynamic capabilities? (RQ3) How does leader-level knowledge contribute to embedding sustainable internationalisation?

Consequently, this paper investigates international leadership as a strategic microfoundation for DCs, operationalised through experiential, judgement-based, and contextual intelligent knowledge. The study contributes to the literature by demonstrating why corporate models cannot be adopted verbatim in academia without accounting for the specific, multidimensional intelligent knowledge of its academic leaders managing international operations.

2 Literature Review

This section establishes the theoretical foundations of the study. It first explores the macro-level DCs framework and its microfoundations within higher education. It then examines the boundary-spanning role of international academic leaders and the cognitive mechanisms underlying their intelligent knowledge. In addition, the section discusses how institutional context shapes organisational resilience, ultimately integrating these perspectives into a conceptual framework that links individual-level cognition to organisational adaptation.

2.1 Dynamic capabilities and their microfoundations in higher education

The DC framework explains how organisations achieve competitive advantage (Teece, Pisano and Shuen, 1997) through three processes: *sensing*, *seizing and transforming* (Teece, Peteraf and Leih, 2016). Recent literature recognises that DCs are rooted in human action, specifically managerial cognition (Helfat and Peteraf, 2015). However, this individual-level expertise must be systematically embedded into collective organisational structures and routines to generate lasting DCs (Felin et al., 2012). This embedding process fundamentally relies on the continuous conversion of individual, tacit expertise into explicit organisational assets (Nonaka and Takeuchi, 1995).

While this theoretical model originated and has been empirically investigated predominantly in the profit-oriented corporate sector (Teece, Peteraf and Leih, 2016), its application to higher education institutions (HEIs) requires significant contextual adaptation. Universities operate under unique governance structures, strict state or foundation regulations, and deeply entrenched organisational histories. In higher education, recent research suggests that dynamic-capability-consistent governance and resource allocation can help institutions respond to policy shifts and financial pressures (Heaton, Teece and Agronin, 2023), while evidence from the COVID-19 context indicates that DCs can also support HEIs in coping with external shocks (Salam et al., 2023). However, applying corporate DC logic encounters a sector-specific tension: massification and economies of scale (Marginson, 2016) create rigidities that conflict with academic quality and institutional values (Noui, 2020). Addressing this tension between scale and agility requires context-sensitive leadership.

2.2 International academic leadership and intelligent knowledge

International academic leaders occupy critical boundary-spanning positions (Whitchurch, 2008), requiring continuous navigation across cultural and organisational domains. The synergy between entrepreneurial and enabling leadership acts as a catalyst for innovation (Beresford-Dey, Howden and Martindale, 2024), which, in the context of HEIs, allows leaders to sense global signals, seize strategic actions, and transform them into internal routines.

Despite leadership's recognised importance, the mechanisms connecting managerial cognition to organisational outcomes remain under-researched (Altintas, 2023). While knowledge management is linked to academic performance (Paudel, Bhattarai and Chalise, 2023), exactly how managerial cognition translates into organisational action and agility requires further exploration. This study builds on Schiuma's (2011) knowledge assets dynamics framework, which emphasises three fundamental management processes: the identification and measurement of knowledge assets, the mapping of their relationships, and the management of knowledge asset flows. While intelligent knowledge is often discussed in relation to technological and AI-driven systems, this study focuses on its human cognitive microfoundations. To conceptualise the micro-level of international academic leadership, the paper integrates Schiuma's (2011) process-oriented logic with Davenport and Prusak's (1998) foundational dimensions of knowledge. In this context, *intelligent knowledge* refers to the experiential, judgement-based, and contextual assets through which leaders interpret signals, make strategic judgements,

and translate external demands into change. How these specific knowledge domains act as the micro-level drivers of institutional agility will be further conceptualised in Section 2.4.

2.3. Institutional context and sustainable internationalisation through organisational resilience

The capacity to operationalise DCs is constrained by the macro-level institutional context. State-funded HEIs face bureaucratic complexity, while private institutions often exhibit greater market-driven flexibility (Levy, 2006). In the context of this study, sustainable internationalisation is conceptualised not merely in terms of maintaining international metrics, but as the institution's ongoing adaptive capacity. From a DCs perspective, sustainable advantage in highly volatile environments depends on organisational resilience (Leih and Teece, 2016; Teece, 2018). Accordingly, sustainable internationalisation represents a university's ability to reconfigure resources and operating models (transforming) in response to exogenous shocks. In this sense, the microfoundation of this resilience lies in the intelligent knowledge of boundary-spanning leaders, through which temporary internationalisation efforts become embedded in formal institutional routines.

2.4 Conceptual framework

This study proposes a conceptual framework (Figure 1) to illustrate how boundary-spanning leaders drive sustainable internationalisation. It synthesises the DCs framework (Teece, 2007; Teece, Peteraf and Leih, 2016) with the microfoundations of knowledge assets (Schiuma, 2011) and the domains of human expertise (Davenport and Prusak, 1998) within the specific institutional context of higher education through a five-step sequence: 1) international academic leadership serves as starting point, (2) the macro-level institutional context defines the boundary conditions, (3) the leader's intelligent knowledge operates as the mediating mechanism, (4) DCs provide the process through which adaptation occurs, and (5) sustainable internationalisation emerges as the organisational outcome.

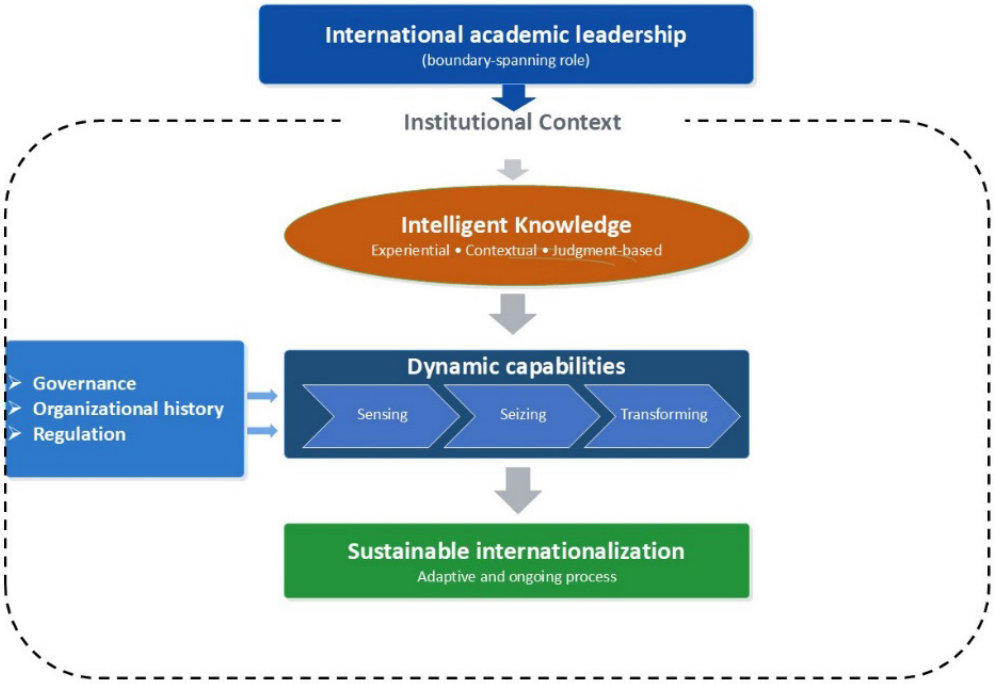


Figure 1: The proposed conceptual framework
Source: Author's own elaboration based on Teece, Peteraf and Leih (2016), Schiuma (2011), and Davenport and Prusak (1998)

To operationalise this theoretical bridge, the present study conceptually adapts Davenport and Prusak's (1998) knowledge dimensions to the specific context of international academic leadership, positioning them as the primary micro-level drivers of institutional agility. Consequently, these domains predominantly inform the sequential phases of DCs (Teece, Peteraf and Leih, 2016) as follows:

- *Experiential knowledge* provides process-based expertise for sensing early signals and scaling operations.

- *Judgement-based* knowledge enables strategic decision-making and tension-resolution (e.g., revenue vs. quality) during the seizing phase.
- *Contextual knowledge* enables boundary-spanning leaders to translate global trends into specific institutional settings, reducing resistance during the transforming phase.

As illustrated in Figure 1, these mapped theoretical intersections do not operate in isolation. The model captures how distinct knowledge domains underpin the sequential organisational phases of sensing, seizing, and transforming. At the same time, it shows that this entire micro-level mechanism is embedded within a broader macro-level institutional context (e.g., state vs. private governance), which determines the conditions under which organisational resilience and sustainable internationalisation can emerge. While these knowledge domains interact fluidly in practice, the framework highlights their primary theoretical intersections with the DC phases.

3 Methodology

3.1 Research design and case selection

This exploratory study uses a comparative multiple-case design for theory elaboration (Yin, 2018). Cases enable *theoretical replication* to examine the conceptual framework from different contextual angles (Eisenhardt, 1989; Eisenhardt and Graebner, 2007; Yin, 2018). Using purposeful sampling for maximum variation (Patton, 2015), three distinct Hungarian HEIs were selected (Table 1). This variation allows for a nuanced exploration of how institutional legacy and governance influence DCs. *Elite interviews* (Harvey, 2011) with boundary-spanning senior international leaders provided the strategic oversight necessary to navigate global trends and internal constraints.

Table 1**:** Profiles of the selected HEIs and interviewees

Case code	Institutional governance	Institutional history	Size	Interviewee role
TS (traditional state)	state-funded, decentralised, research focus	>100 years	20,000+ students	Vice-Rector for International Affairs
CS (centralised state)	state-funded, centralised	10+ years	6,000+ students	Vice-Rector for International Affairs
PM (private market)	British-accredited private institution	30+ years	1,000+ students	Rector

Note: In the PM case, the Rector fulfils the boundary-spanning role due to a flatter hierarchy.

3.2 Data collection and analysis

Data were gathered through semi-structured elite interviews. An eight-question guide, mapped to the sensing, seizing, and transforming phases, used probing questions to elicit concrete strategic episodes (Kvale and Brinkmann, 2009). Interviews were conducted in Hungarian to preserve deep contextual nuance (van Nes et al., 2010), then transcribed and selected representative quotes were translated into English.

Deductive thematic analysis (Braun and Clarke, 2006) categorised data into an Excel matrix based on the conceptual framework (context, knowledge dimensions, DC phases). Analysis followed a two-step logic (Miles, Huberman and Saldaña, 2014): a within-case analysis of each HEI, followed by a cross-case comparison to identify patterns and variances across governance models.

4 Findings

Analysis reveals that while all three institutions employ DCs to navigate internationalisation, micro-level leadership knowledge and macro-level context produce distinct trajectories. The synthesis follows the conceptual model's logic (Figure 1). The analysis first explores how institutional context creates inertia or fosters agility, before examining the activation of leaders' intelligent knowledge and the resulting tensions between massification and quality. This leads to an operational overview of the sensing, seizing, and transforming phases, ultimately identifying distinct pathways to organisational resilience.

4.1. The role of institutional context: inertia versus agility

The institutional environment fundamentally impacts internationalisation's speed and governance. For the TS university, historical legacy provides reputational capital but immense organisational inertia. As the TS leader

noted, faculties are "highly sceptical of centralisation goals," which transforms any strategic shift into a "pull-back force" requiring extensive internal negotiations.

Conversely, the CS university's hierarchical framework necessitates "highly central management," which allows the institution to "react faster to exogenous shocks because it is more centralised." However, the most extreme agility was observed in the PM institution, where being unencumbered by bureaucracy enabled rapid decision-making. The PM rector explained this operational freedom through stark financial realities: in its early stages, the institution was "on the verge of bankruptcy", necessitating radical, immediate market pivots over prolonged academic consensus-building.

4.2. The microfoundations: activating intelligent knowledge

The microfoundations driving the internationalization process, conceptualised as leaders' intelligent knowledge of boundary-spanning leaders, varied by institutional type. The TS leader utilised deep contextual knowledge to overcome internal resistance. Having previously served at the faculty level, the leader gained legitimacy because the apparatus "knew that I had suffered with the central administration just as they did". Similarly, the CS leader mobilised experiential knowledge. Despite serving as a vice-dean, like their TS counterpart, the CS leader emphasised a bottom-up trajectory, having started as an international coordinator. This meant they already "knew the international processes well" from the ground up, allowing them to scale this practical, operational expertise directly to the rectorate level as a "huge resource".

Conversely, PM microfoundations were rooted in business judgement over traditional academic socialisation. The PM leader emphasised that survival required "looking for business solutions, not academic ones." Lacking a regional blueprint, the PM's international strategy became a "trial-and-error game," relying heavily on the leader's entrepreneurial risk-taking rather than established institutional data.

4.3. Navigating the core tension: massification versus quality

Leaders navigated the tension between revenue-driven massification and academic quality through distinct strategies. TS employed a structural compromise: fee-paying bachelor's students generate revenue, while doctoral students "bring the necessary energy from a scientific output perspective" to sustain quality.

The CS leader criticised ranking pressures, noting that pursuing quotas ("let international students come, because it pushes up the ranking") can result in admitting candidates who "hinder the academic progression of domestic students." The PM rector articulated a private paradox: "Until you have an established brand name, you sacrifice quality... because you must generate a return for the owners." To mitigate this aggressive massification, the PM institution strategically utilised its British accreditation as a strict quality assurance shield to legitimise degrees on the global market. However, the PM leader noted that having finally achieved stability, the institution is now positioned to initiate a very gradual, deliberate shift toward enhancing student quality.

4.4. Operationalising dynamic capabilities

All institutions showed strong sensing: TS utilised embassy networks, CS monitored conferences, and PM recognised the marketing value of European visas. However, seizing and transforming diverged significantly.

During seizing, TS university relied on incremental experimentation, testing "small variations" to preserve "fragile faculty consensus". Conversely, the CS leader utilised top-down authority for strategic commitments, acknowledging that seizing international markets "requires financial sacrifices from the universities." The PM leader, driven by financial pressure, engaged in high-stakes risk-taking, describing seizing as a forced "trial-and-error game."

The most drastic differences emerged in the transforming phase. At TS, administrative complexity delays change: "you have to get 100 or 200 people to change their mindset, and obviously, that is much slower." The CS institution managed to embed certain processes through formalisation (e.g., "written rector's directives remained"), yet informal initiatives collapsed post-leadership. In contrast, the PM university exhibited radical transformation capabilities. During the pandemic, it executed a "25% staff layoff" and, in a previous instance, rapidly phased out unprofitable language programmes: drastic, immediate reconfigurations that, as the leader noted, "would never work in a state institution".

4.5. Pathways to resilience

As conceptualised (see Section 2.4), this study operationalises sustainable internationalisation through organisational resilience. The empirical data reveal that HEIs achieve this resilient state through entirely different pathways.

For the TS university, resilience stems from its massive size and structural heterogeneity: its faculties function as “small laboratories where you can test mini worlds” and cross-pollinate best practices. For the CS university, resilience is secured through strategic international embeddedness. The leader identified prestigious European grants as the ultimate safety net, as these 5-to-10-year commitments “strongly anchor the institution into the European Higher Education Area”, providing a buffer against domestic financial shocks. Finally, the private PM university’s resilience is derived from two distinct factors. First, its predominantly non-academic leadership mindset allows the institution to immediately reframe and resolve external shocks as pragmatic business problems. Second, resilience is rooted in a highly autonomous, “live and let live” organisational culture. This emphasis on personal freedom generates remarkable personnel stability and low staff turnover, acting as the institution’s primary shield during turbulent times.

Collectively, these divergent micro-level solutions suggest that there is no universal template: the sustainability of internationalisation depends entirely on the intelligent alignment of leadership knowledge with the specific institutional context.

5 Discussion

This study unpacks the black box of DCs in higher education by examining the microfoundations of international academic leadership. This section addresses the three research questions by integrating empirical findings with the theoretical frameworks of dynamic capabilities (Teece, 2007; Teece, Peteraf and Leih, 2016) and synthesising Schiuma’s (2011) knowledge asset logic with Davenport and Prusak’s (1998) dimensions. By operationalising intelligent knowledge as a multidimensional cognitive asset, the following analysis demonstrates how individual-level leadership drives institutional adaptation within the international higher education landscape.

5.1. The micro-macro linkage: activating dynamic capabilities

In response to RQ1 (how leaders apply intelligent knowledge across DC phases) and RQ2 (how institutional context shapes this operationalisation), the cross-case analysis suggests that corporate DC models require significant contextual adaptation for HEIs. The findings corroborate Helfat and Peteraf’s (2015) argument that managerial cognition constitutes a key microfoundation of DCs and extend Altintas’s (2023) framework by illustrating how individual-level cognitive processes drive organisational knowledge management. This transformation aligns with the knowledge-creation logic proposed by Nonaka and Takeuchi (1995), suggesting that international leaders’ intelligent knowledge is systematically converted from individual tacit insights into explicit institutional routines.

During sensing, all leaders relied heavily on experiential knowledge (Davenport and Prusak, 1998) to identify global weak signals. However, the institutional context (RQ2) dictated the focus of this sensing. In the PM case, operating with greater autonomy from state bureaucracy, consistent with established observations on private institutional agility and market-driven adaptation (Levy, 2006), experiential knowledge facilitated the rapid identification of revenue-generating market niches. Conversely, in the TS and CS cases, sensing was shaped by state expectations, requiring leaders to scan for signals that corresponded to national regulations and traditional academic prestige.

The critical divergence emerged during seizing, where judgement-based knowledge (Davenport and Prusak, 1998) assumed particular importance. While PM seizing mirrored corporate logic via rapid resource mobilisation (Teece, Peteraf and Leih, 2016), leaders at TS and CS institutions faced considerable organisational inertia. For these boundary-spanning executives (Whitchurch, 2008), seizing required careful balancing, utilizing intuitive judgement to reconcile systemic pressure for revenue growth with entrenched academic values. In these state environments, judgement-based knowledge acted as the crucial mechanism for navigating institutional hesitancy and mitigating deeply rooted faculty resistance.

Finally, the transforming phase highlighted the importance of contextual knowledge (Davenport and Prusak, 1998) in addressing internal resistance. The PM institution exhibited high organisational agility, allowing leadership to implement structural changes rapidly. However, in TS and CS environments, the legacy of academic silos presented significant barriers. Leaders achieved strategic transformation by reconfiguring resources and operating models (Teece, 2018). To accomplish this, they actively managed knowledge asset flows (Schiuma, 2011) and embedded individual decisions into collective routines (Felin et al., 2012), translating market demands into acceptable academic discourse to mitigate faculty pushback.

5.2. Resolving the massification tension and building resilience

RQ3 explored how leader-level intelligent knowledge contributes to the long-term embedding of sustainable internationalisation. Given the profound tension between volume-based massification and academic quality (Marginson, 2016; Noui, 2020), findings show that boundary-spanning leaders use this cognitive asset to navigate this conflict, achieving a form of organisational ambidexterity and thereby generating organisational resilience.

Sustainable internationalisation manifests as organisational resilience, the continuous adaptive capacity to survive exogenous shocks (Salam et al., 2023). The pathways to this resilience differ fundamentally based on macro-level governance (Heaton, Teece and Agronin, 2023). The TS university achieves resilience through its massive scale and decentralised faculty networks, which act as "mini laboratories" for international practices. The CS university builds resilience through deep structural embeddedness in European grant networks. Meanwhile, the PM institution achieves resilience through raw organisational agility (Teece, Peteraf and Leih, 2016) and a highly cohesive organisational culture that values rapid adaptation over academic traditions. In all three scenarios, it is the international leader's specific intelligent knowledge that helps institutions manage the pressures generated by massification, enhancing their evolutionary fitness and transforming temporary crisis management into embedded institutional resilience (Leih and Teece, 2016).

5.3. Theoretical contributions

This study makes two primary theoretical contributions. First, it extends the DCs framework (Teece, 2007) in the higher education context by integrating Schiuma's (2011) process-oriented logic with Davenport and Prusak's (1998) knowledge dimensions, conceptualised here as intelligent knowledge, as a necessary microfoundation. It demonstrates that the sequence of sensing, seizing, and transforming (Teece, Peteraf and Leih, 2016) in academia is not a uniform, macro-level routine, but an individual-level process driven by the cognitive domains of boundary-spanning international academic leaders. Second, the study contributes to the internationalisation literature by demonstrating that corporate agility models are inherently constrained by the macro-level governance structures of HEIs (state-funded versus private).

6 Conclusions

As this study has demonstrated, achieving sustainable internationalisation in today's turbulent higher education sector goes beyond mere strategic intent: it fundamentally depends on the dynamic capabilities embedded within institutional leadership. This paper argues that navigating the inherent tension between volume-driven massification and academic quality requires well-developed dynamic capabilities (DCs). Through a comparative analysis of three Hungarian HEIs, the study concludes that sustainable internationalisation is not achieved through rigid, macro-level corporate models, but through the micro-level intelligent knowledge of boundary-spanning leaders responsible for internationalisation. It is the leader's ability to align experiential, judgement-based, and contextual knowledge with the specific governance structure of the institution that drives sensing, seizing, and transforming activities, effectively generating organisational resilience.

From a practical and managerial perspective, the findings offer critical implications for university leadership, revealing that structural and administrative changes alone are insufficient without the right cognitive microfoundations. Vice-rectors must actively cultivate all three domains of intelligent knowledge: experiential knowledge to sense global signals and scale operations, judgement-based knowledge to make bold decisions under high uncertainty during the seizing phase, and contextual knowledge to overcome institutional inertia during transformation. In addition, to prevent informal initiatives from collapsing during leadership succession, universities must systematically embed this individual expertise into formal institutional routines, ensuring long-term knowledge retention and organisational resilience.

While drawn from the Hungarian higher education context, the implications extend far beyond this geographical boundary. The findings are highly relevant for a broad segment of HEIs globally that operate outside the elite tier. For such institutions, leadership's cognitive agility remains a key source of adaptive capacity. Future research should quantitatively test the proposed conceptual framework across diverse national contexts and explore how mid-level academic managers, such as faculty deans, interact with top leadership to execute DCs and drive institutional transformation in the field of internationalisation.

7 Ethics Declaration

Formal ethical clearance was not required for this study. The author declares that the research was conducted in accordance with standard academic ethical guidelines. Informed consent was obtained from all interview participants, and institutional anonymity has been strictly maintained to ensure confidentiality.

8 AI Declaration

The author declares that she did not use any generative artificial intelligence tools in the preparation or editing of this document. All information and ideas drawn from other sources are properly cited and referenced according to the guidelines of the IFKAD Principles and Policies for AI use.

References

Altintas, G. (2023) "Microfoundations of Knowledge-Based Dynamic Capabilities", *M@n@gement*, Vol. 26, No. 2, pp 54–67.

Beresford-Dey, M., Howden, S. and Martindale, L. (2024) "Complexity Leadership Theory and its application in higher education: using duoethnography to explore enabling leadership during a time of uncertainty", *International Journal of Leadership in Education*, pp 1–26.

Braun, V. and Clarke, V. (2006) "Using thematic analysis in psychology", *Qualitative Research in Psychology*, Vol. 3, No. 2, pp 77–101.

Davenport, T.H. and Prusak, L. (1998) *Working Knowledge: How Organizations Manage What They Know*, Boston, MA: Harvard Business School Press.

de Wit, H. (2020) "Internationalization of Higher Education: The Need for a More Ethical and Qualitative Approach", *Journal of International Students*, Vol. 10, No. 1, pp i-iv.

Eisenhardt, K.M. (1989) "Building theories from case study research", *Academy of Management Review*, Vol. 14, No. 4, pp 532–550.

Eisenhardt, K.M. and Graebner, M.E. (2007) "Theory building from cases: Opportunities and challenges", *Academy of Management Journal*, Vol. 50, No. 1, pp 25–32.

Felin, T., Foss, N.J., Heimeriks, K.H. and Madsen, T.L. (2012) "Microfoundations of routines and capabilities: Individuals, processes, and structure", *Journal of Management Studies*, Vol. 49, No. 8, pp 1351–1374.

Harvey, W.S. (2011) "Strategies for conducting elite interviews", *Qualitative Research*, Vol. 11, No. 4, pp 431–441.

Heaton, S., Teece, D.J. and Agronin, E. (2023) "Dynamic capabilities and governance: An empirical investigation of financial performance of the higher education sector", *Strategic Management Journal*, Vol. 44, No. 2, pp 520–548.

Helfat, C.E. and Peteraf, M.A. (2015) "Managerial cognitive capabilities and the microfoundations of dynamic capabilities", *Strategic Management Journal*, Vol. 36, No. 6, pp 831–850.

Knight, J. (2004) "Internationalization remodeled: Definition, Approaches, and Rationales", *Journal of Studies in International Education*, Vol. 8, No. 1, pp 5–31.

Kvale, S. and Brinkmann, S. (2009) *InterViews: Learning the Craft of Qualitative Research Interviewing*, 2nd ed., Thousand Oaks, CA: Sage Publications.

Leih, S. and Teece, D.J. (2016) "Campus leadership and the entrepreneurial university: A dynamic capabilities perspective", *Academy of Management Perspectives*, Vol. 30, No. 2, pp 182–210.

Levy, D. C. (2006) "The unanticipated explosion: Private higher education's global surge", *Comparative Education Review*, Vol. 50, No. 2, pp 217–240.

Marginson, S. (2016) "The worldwide trend to high participation higher education: dynamics of social stratification in inclusive systems", *Higher Education*, Vol. 72, No. 4, pp 413–434.

Miles, M.B., Huberman, A.M. and Saldaña, J. (2014) *Qualitative Data Analysis: A Methods Sourcebook*, 3rd ed., Thousand Oaks, CA: Sage Publications.

Nonaka, I. and Takeuchi, H. (1995) *The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation*, New York: Oxford University Press.

Noui, R. (2020) "Higher education between massification and quality", *Higher Education Evaluation and Development*, Vol. 14, No. 2, pp 93–103.

Patton, M.Q. (2015) *Qualitative Research & Evaluation Methods: Integrating Theory and Practice*, 4th ed., Thousand Oaks, CA: Sage Publications.

Paudel, K.P., Bhattarai, P.C. and Chalise, M. (2023) "Interdependencies between knowledge management and academic performance in higher educational institutions", *VINE Journal of Information and Knowledge Management Systems*, Vol. 53, No. 4, pp 748–765.

Salam, A., Hussain, S., Awan, W.A. and Amin, H. (2023) "Leveraging the Power of Open Innovation and Dynamic Capabilities to Recover from COVID-19 Impacts: A Study of Higher Education Institutions of Pakistan", in Sultan, P. (ed.) *Innovation, Leadership and Governance in Higher Education: Perspectives on the Covid-19 Recovery Strategies*, Singapore: Springer, pp 119–138.

Schiuma, G. (2011) "Managing and Measuring Knowledge Assets Dynamics for Business Value Creation in Organisations", in Schiuma, G. (ed.) *Managing Knowledge Assets and Business Value Creation in Organizations: Measures and Dynamics*, Hershey, PA: IGI Global, pp 13–29.

Seeber, M., Cattaneo, M., Huisman, J. and Paleari, S. (2016) "Why do higher education institutions internationalize? An investigation of the multilevel determinants of internationalization rationales", *Higher Education*, Vol. 72, No. 5, pp 685–702.

Teece, D.J. (2007) "Explicating dynamic capabilities: The nature and microfoundations of (sustainable) enterprise performance", *Strategic Management Journal*, Vol. 28, No. 13, pp 1319–1350.

Teece, D.J. (2018) "Business models and dynamic capabilities", *Long Range Planning*, Vol. 51, No. 1, pp 40–49.

Teece, D.J., Peteraf, M. and Leih, S. (2016) "Dynamic capabilities and organizational agility: Risk, uncertainty, and strategy in the innovation economy", *California Management Review*, Vol. 58, No. 4, pp 13–35.

Teece, D. J., Pisano, G. and Shuen, A. (1997) "Dynamic capabilities and strategic management" *Strategic Management Journal*, Vol. 18, No. 7, pp 509–533.

van Nes, F., Abma, T., Jonsson, H. and Deeg, D. (2010) "Language differences in qualitative research: is meaning lost in translation?", *European Journal of Ageing*, Vol. 7, No. 4, pp 313–316.

Whitchurch, C. (2008) "Shifting identities and blurring boundaries: The emergence of third space professionals in UK higher education", *Higher Education Quarterly*, Vol. 62, No. 4, pp 377–396.

Yin, R.K. (2018) *Case Study Research and Applications: Design and Methods*, 6th ed., Thousand Oaks, CA: Sage Publications.

Human–LLM Collaboration in Tactical Production Decision-Making: A Literature Review

Ines Ben Atitallah, Remsei Sándor, Heba F. Zaher
Széchenyi István Egyetem, Győr, Hungry

Abstract

This research seeks to investigate how large language models contribute to decision-making in manufacturing settings, with special emphasis on how LLM use influences decision effectiveness and the creation of a collaborative decision-making process between humans and LLMs.

The study adopts a literature review following PRISMA 2020 guidelines. A total of 145 papers were sourced from the Scopus and Web of Science databases (2020–2026). After duplicate removal and title and abstract analysis, 53 papers remained for detailed review. Out of these, 38 papers were included in the study. An additional 9 sources were selected through manual searches, bringing the total to 47. The studies were analysed across four dimensions: decision context, LLM involvement, human–LLM collaboration, and decision outcomes.

The results reveal that LLMs are primarily utilised for decision support purposes at the tactical level in production environments, such as planning, scheduling, diagnostics, and coordination. Despite demonstrating significant abilities in data processing and reasoning, the effect of LLMs on decision quality remains uncertain. Current literature mainly emphasises the technical aspects of LLMs rather than the human-machine interaction process, trust-building mechanism, and integration with organisational practices. Studies usually employ a human-in-the-loop framework, where decision-making remains in human hands.

This study focuses on peer-reviewed journal articles published in English and indexed in Scopus and Web of Science. As a result, some relevant research from other sources or languages may not be included. The study offers insights for organisations adopting LLM-based decision-support systems, emphasising the need for transparency, human intervention, and alignment with decision-making contexts and processes. This study provides a clear picture of how current research on LLMs in manufacturing decision-making is structured and highlights the main gaps that still exist. It shows that future research should go beyond focusing solely on technical capabilities and pay more attention to how humans actually interact with LLMs, as well as to how these systems influence decision quality in real-world situations.

The paper makes an original contribution by integrating scattered literature and highlighting significant knowledge gaps in decision-quality evaluation and human-LLM cooperation in tactical production decision-making.

Keywords – Large Language Models; Tactical Decision-Making; Manufacturing; Human–AI Collaboration; Decision Support Systems

Paper type – Literature review

1 Introduction

The rapid development of AI, particularly large language models (LLMs), has significantly expanded decision support capabilities. LLMs can significantly enhance decision-making in manufacturing and production. Specifically, because tactical decisions in manufacturing are complex and require substantial information, LLMs can provide decision-makers with valuable insights. At the same time, the current scholarly literature is fragmented in its discussion of the application of LLMs in decision-making. In other words, most studies focus on the technical capabilities of LLMs (optimisation, reasoning, and system integration). However, few articles discuss how LLMs interact with humans, what role they play in decision-making, and how they can be successfully implemented in production processes.

Considering the above gap, it is necessary to understand the influence large language models have on decisions made in manufacturing companies. The objective of the present paper is to examine the influence of artificial intelligence technologies on tactical decision-making and design processes with the aim of integrating human and artificial intelligence decisions. An analysis of prior research on this issue will be conducted to assess the current situation and identify knowledge gaps.

2 Review Protocol

A structured review protocol was developed to define the scope, research questions, and procedures guiding the literature review. While this study does not follow a fully systematic literature review (SLR) approach,

it draws on selected principles from established guidelines, such as PRISMA, to enhance transparency, consistency, and methodological rigor in the review process.

The review is guided by the following research questions:

- ***RQ1:*** *How can organisations design processes and systems to effectively integrate human judgment with LLM-based decision support?*
- ***RQ2:*** *How does LLM-supported decision-making influence the quality of tactical production decisions?*

The above queries guided the entire review process, from searching to the selection of studies. In this regard, this review adhered to the PRISMA 2020 guidelines (Page et al., 2021) in the presentation of the search and selection process and was done in adherence to the systematic review methods provided by Kitchenham and Charters (2007).

Figure 1 provides a flow diagram showing the process through which the studies were selected. The Rayyan tool was used in the handling of records and decisions during the screening process.

2.1 Search strategy

A comprehensive, automated search was conducted in Scopus and Web of Science to capture peer-reviewed studies (2020–2026) on LLMs and manufacturing decision-making. The core Boolean query was:

("decision making" OR "decision-making") AND ("large language model" OR "LLM" OR "LLMs") AND ("manufactur*" OR "production")*

This query was applied identically in both databases. Filters were applied to limit results to English-language journal articles from 2020 to 2026, focusing on subject categories relevant to business, management, decision sciences, engineering, and operations research. The searches were last run on 30-03-2026. Initial results totalled 145 records: 74 from Scopus and 71 from Web of Science (Table 1). All retrieved citations were exported (RIS format) and imported into Rayyan for management and screening.

Table 1. Search log detailing databases, queries, filters, and the initial number of articles (2020–2026).

Database	Search String	Filters	Number of articles
Scopus	("decision making" OR "decision-making") AND ("large language model*" OR "LLM" OR "LLMs") AND ("manufactur*" OR "production")	2020–2026; Decision Sciences; Business & Management; Social Sciences; Multidisciplinary	74
Web of Science		2020–2026; Business; Management; Economics; Engineering Manufacturing; Operations Research; Industrial Engineering	71
Total			145

2.2 Study selection process (PRISMA)

The screen adopted the PRISMA 2020 protocol (Matthew J. Page et al., 2021) in four distinct phases: identification, screening, eligibility, and inclusion. Articles were eligible for inclusion if they were published during the years 2020–2026, written in English, and dealt specifically with LLMs or other AI technologies used in decision-making processes in manufacturing/production environments, focusing primarily on human-AI decision interactions and their decision quality effects. Exclusion criteria comprised studies outside the manufacturing domain (e.g., healthcare or finance), papers not involving decision support, non-academic sources (e.g., white papers), and articles lacking sufficient methodological or contextual detail. These criteria were consistently applied across all screening stages to ensure alignment with the research questions.

After importing 145 articles in the Rayyan database, 24 duplicates were filtered out, leaving 121 unique articles. Screening was then performed based on titles and abstracts, and 68 articles were excluded since they were not relevant. Thus, a total of 53 papers underwent full-text screening. Out of the 53 articles, 15 were excluded for being less relevant or having inadequate methodology. To further increase the scope of the study, an additional nine papers that were found during the early stages of background research before keyword search were included in the review. Thereby, the final number of articles in the review is 47. Table 2 summarises these results, and Figure 1 illustrates the selection process.

Table 2. PRISMA study selection summary (2020–2026).

PRISMA Stage	Number of Records
Records identified (databases)	145
Duplicates removed	24
Records screened	121
Records excluded (title/abstract)	68
Full-text articles assessed	53
Full-text articles excluded	15
Additional records (manual search)	9
Final included studies	47

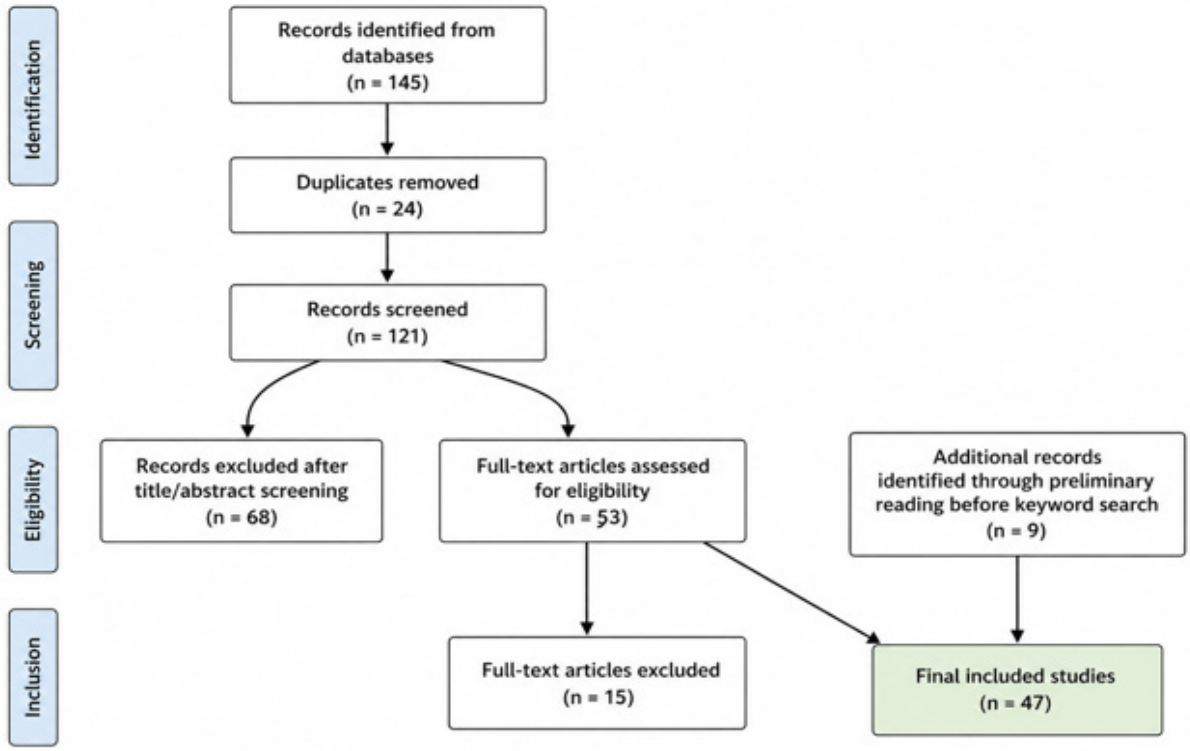


Figure 1. PRISMA 2020 flow diagram of study selection

2.3 Data extraction

Data were extracted from each study into a standardised form. Key fields included: author(s) and year; publication source; type of LLM or AI tool evaluated; decision context (e.g., scheduling, resource allocation, maintenance); decision level (tactical, operational); nature of human–AI interaction (e.g., advisory, collaborative, or autonomous); evaluation methodology (e.g., case study, experiment, simulation); and reported performance outcomes (e.g., efficiency, accuracy, timeliness). Further, data on how the human factor was defined in each study, the level of interaction between the human and the LLM, and any discussion of decision-making were collected. When available, the effects on decision-quality outcomes were noted; however, since very few studies measured decision quality, proxies for decision quality (such as improved performance or decreased errors) were also included. Any process or system design suggestions for integrating the human and the LLM were also collected. The selection of the fields was guided by RQ1 (decision-quality outcomes) and RQ2 (process and system design factors).

3 Overview of the Reviewed Studies

To provide a comprehensive understanding of the literature, this section will present a discussion of all 47 reviewed articles from two perspectives: concept classification and factor analysis of decision-making processes.

In the first case, a concept matrix (Webster & Watson, 2002) will be used in order to examine the current state of the literature. In the second one, a factor analysis will be applied to examine the collaborative process between humans and LLMs during production decisions.

3.1 Concept matrix of LLM applications in manufacturing

Table 3 shows the classification of reviewed studies using a concept matrix approach. The literature shows a strong concentration on LLM decision support, indicating that most research focuses on improving technological performance in terms of efficiency, optimisation, and automation in production decision-making.

Table 3. Classification of reviewed studies using a concept matrix approach.

Concept	Description (Focus)	Representative Studies
LLM decision support	General use of LLMs as decision-support tools in production and manufacturing systems, focusing on efficiency, optimisation, monitoring, and assistance	Vafeiadis et al. (2025); Gholami (2024); Fan et al. (2026); Cimino et al. (2026); Li et al. (2025, 2026); Lin et al. (2025); Tong et al. (2026); Liu et al. (2026); Wu et al. (2026); Sadeghi et al. (2026); Garcia et al. (2024); Colabianchi et al. (2024)
Decision support in scheduling	Application of LLMs in tasks such as sequencing, production scheduling, coordination, and real-time adjustment of activities under constraints	Genetti et al. (2026); Wu et al. (2025); Yang et al. (2026)
LLM-assisted diagnosis	Use of LLMs for fault detection, maintenance, troubleshooting, and diagnostic reasoning in industrial systems	Marandi et al. (2025); Nie et al. (2026); Wu et al. (2025); Jadhav et al. (2025)
Knowledge-driven decisions	Use of LLMs to extract, structure, and apply domain knowledge (e.g., tacit knowledge, ontologies, knowledge graphs) in decision processes	Rank et al. (2025); Ioshchikhes et al. (2026); Skrzek et al. (2026); Yan et al. (2026)
Planning decision support	Use of LLMs in production planning, system design, and resource allocation across multiple stages, often involving forecasting and optimisation	Werheid et al. (2025); Huang et al. (2026); Tan et al. (2026); Teng et al. (2025); Cimino et al. (2026)
Human–LLM collaboration	Studies focusing on interaction, co-pilot systems, trust, explainability, and human-in-the-loop decision-making	Colabianchi et al. (2024); Sridhar et al. (2026); Favier et al. (2025); Wang et al. (2026); Cai et al. (2026); Sabbah & Li (2025); Elgeddawy (2026); Passerini et al. (2025)
Human–robot–LLM interaction	Integration of LLMs into robotic or embodied systems where humans provide guidance, perception, or commands	Liu et al. (2026 – AR-assisted motion planning); Ding et al. (2026); Dong et al. (2025)
AI autonomy in decisions	Fully or highly autonomous LLM-driven decision-making systems with minimal or no human involvement	Wu et al. (2025); Yang et al. (2026); Ajraou et al. (2026)

More specific applications, such as scheduling, diagnosis, and planning, are present but less common, suggesting that research is fragmented across individual production functions rather than addressing decision-making as an integrated process. Although some studies explore human–LLM interaction, this remains one of the least developed areas. While interest is increasing, few studies examine interaction processes, trust-building, or decision quality in depth. At the same time, another trend is evident toward increasing automation, with several studies focusing on autonomous systems (e.g., robot-based manufacturing), where human involvement becomes limited.

3.2 Key factors influencing human–LLM collaboration

In the previous section, the concept matrix illustrates the research that has been conducted. However, this section further investigates the literature on human–LLM interaction to identify the factors influencing LLM-based decision-making in production environments.

The findings demonstrate that decision-making effectiveness is influenced by several dimensions, including decision attributes, LLM features, human factors, organisational settings, and the production context.

Table 4. Factors shaping LLM-supported decision-making.

Category	Key Factor	Short Explanation	Supporting Studies
Decision Characteristics	Decision complexity and uncertainty	LLMs are more useful in complex and uncertain decisions but also increase the risks of misinterpretation.	Cai et al. (2026); Wu et al. (2026); Tong et al. (2026); Sadeghi et al. (2026); Yang et al. (2026); Garcia et al. (2024)
	Time constraints (real-time vs planned)	Real-time decisions require speed, while planned decisions allow validation, affecting decision reliability.	Yang et al. (2026); Cai et al. (2026); Jadhav et al. (2025); Werheid et al. (2025); Fan et al. (2026)
	Decision structure	LLMs are commonly applied in semi-structured or unstructured tasks, where human interpretation remains important.	Sadeghi et al. (2026); Sabbah & Li (2025); Passerini et al. (2025); Favier et al. (2025)
LLM Characteristics	Explainability of outputs	Clear explanations improve trust and support better decision validation.	Passerini et al. (2025); Sridhar et al. (2026); Elgeddawy (2026); Cai et al. (2026); Genetti et al. (2026)
	Reasoning capability	Strong reasoning improves adaptability, but its effectiveness may vary depending on context.	Liu et al. (2026); Jadhav et al. (2025); Fan et al. (2026); Favier et al. (2025)
	Reliability	Inconsistent or hallucinated outputs can reduce decision quality.	Wang et al. (2026); Cai et al. (2026); Passerini et al. (2025); Cariaga (2025); Garcia et al. (2024)
	System integration	Integration with systems (KG, DT, agents) improves accuracy and usefulness.	Wu et al. (2024); Wu et al. (2026); Teng et al. (2025); Yang et al. (2026); Garcia et al. (2024)
Human Factors	Trust and reliance	Trust affects how much users depend on LLM outputs.	Elgeddawy (2026); Sabbah & Li (2025); Wang et al. (2026); Cariaga (2025); Genetti et al. (2026)
	Expertise	Skilled users better interpret and validate LLM outputs.	Liu et al. (2026); Passerini et al. (2025); Rank et al. (2025)
	Cognitive load and usability	LLMs can reduce effort, but poor design may increase confusion.	Colabianchi et al. (2024); Werheid et al. (2025)
	Decision authority	Humans usually retain final control in decision-making.	Cai et al. (2026); Elgeddawy (2026); Sridhar et al. (2026); Werheid et al. (2025)
Organisational Factors	Workflow integration	Effective integration into workflows improves decision support.	Colabianchi et al. (2024); Garcia et al. (2024); Werheid et al. (2025)
	Adoption culture and governance	Organizational readiness and policies influence adoption.	Elgeddawy (2026); Cariaga (2025); Wulf et al. (2024)
	Training and change management	Training is needed to ensure effective use of LLMs.	Passerini et al. (2025); Wang et al. (2026); Cariaga (2025)
Production Context	Industry/domain	LLM effectiveness varies across industries and applications.	Wu et al. (2026); Sadeghi et al. (2026); Tong et al. (2026); Garcia et al. (2024)
	Environment dynamics	Dynamic environments require higher adaptability from LLMs.	Wu et al. (2026); Yang et al. (2026); Cai et al. (2026); Liu et al. (2026)
	Data availability and quality	High-quality data is essential for reliable LLM outputs.	Wu et al. (2024); Garcia et al. (2024); Wu et al. (2024 – ProcessCarbonAgent)
	Task type	LLM performance varies depending on the type of decision task.	Colabianchi et al. (2024); Jadhav et al. (2025); Liu et al. (2026)

From the analysis in Table 4 above, it becomes clear that apart from being influenced by the technology itself, decision making supported by LLM technology is affected by many interrelated factors. Even though certain features of the technology, such as reasoning capability, explainability, and system integration, are significant, their impact on overall efficiency depends heavily on how they relate to human decision-makers and the context. Specifically, human variables such as expertise and decision-making ability significantly influence how outputs from an LLM are understood and utilised. At the same time, organisational factors such as

integration into workflow, governance and training play an important role in determining the success of implementation. Moreover, production variables contribute, including task types, data quality, and a dynamic environment.

4 Human–LLM Collaboration in Tactical Production Decision-Making: Roles, Patterns, and Challenges

This section analyses how LLMs are applied in tactical production decision-making, by analysing their functions, relations with human decision makers, and major obstacles to their application.

4.1. Decision types and context

Most of the decision-making processes aided by LLMs fall into the category of semi-structured decisions, where models suggest recommendations and humans make the final decisions. This can be seen in fields such as scheduling and planning, where the outputs need to reviewed and adjusted. These contexts are typically complex, involving multiple constraints and dependencies, and often require real-time or short-term responses.

Overall, LLM-supported decisions are primarily semi-structured, complex, and time-sensitive, reflecting the nature of tactical production decision-making. However, limited attention has been given to understanding human–LLM interaction in these settings (Genetti et al., 2026; Wu et al., 2025; Yang et al., 2026; Tan et al., 2026).

4.2 LLM in tactical production decision-making

LLMs are primarily used for tactical production tasks, such as scheduling and coordination, production planning, and maintenance or diagnostics (Genetti et al., 2026; Wu et al., 2025; Tan et al., 2026; Marandi et al., 2025). Additional uses include monitoring, robotic coordination, and sustainability-related decisions (Ajraou et al., 2026; Liu et al., 2026).

These applications span shop-floor activities (e.g., assembly, machining) and system-level contexts (e.g., supply chains, energy systems). Overall, LLMs are most often used in information-rich, dynamic, and semi-structured environments requiring timely decisionsHuang et al., 2026; Werheid et al., 2025; Wu et al., 2026).

4.3 Human–LLM collaboration patterns

Three main types of collaboration between humans and LLMs can be identified. First, fully autonomous decision-making by LLMs without human involvement is rare, though it occurs in some highly automated systems, such as multi-agent production scheduling or robotic control (e.g., Wu et al., 2024). Second, a common approach is human validation, where LLMs generate suggestions that are then reviewed and confirmed by human decision-makers, for example, in fault diagnosis or maintenance support systems (e.g., Nie et al., 2026; Sridhar et al., 2026). Third, some studies describe a more interactive form of collaboration, where humans and LLMs work together through continuous communication and joint decision-making, such as in co-pilot systems or human–robot collaboration environments (e.g., Ding et al., 2026; Sabbah & Li, 2025; Colabianchi et al., 2024). Overall, most research focuses on the last two approaches, where human involvement remains essential.

4.4 LLM Capabilities and limitations

The literature highlights that LLMs are effective in processing large volumes of data and supporting reasoning tasks such as planning and diagnosis (e.g., Nie et al., 2026; Liu et al., 2026). They can also generate context-aware explanations, improving accessibility for users (Genetti et al., 2026). Nevertheless, this technique has important constraints. Explainability is low, which complicates model reasoning, and reliability problems, such as hallucinations and inconsistency, undermine trust in LLM recommendations. (Wang et al., 2026; Passerini et al., 2025).

4.5 Challenges in real production contexts

Using LLMs in real production environments comes with several practical challenges. One of the main issues is trust, as users are often cautious when relying on model outputs that may not always be fully transparent or consistent (Dong et al., 2025; Wang et al., 2026). Usability is another important factor, since these systems need to be easy to understand and interact with in order to support decision-making effectively. At the same time, integrating LLMs into existing industrial environments is not straightforward. It requires aligning them with

current systems, workflows, and organisational processes, which can be complex and time-consuming (Fan et al., 2024; Elgeddawy, 2026).

5 Discussion

In examining the selected articles, the following consistent trends in the use of LLMs in production decision-making can be highlighted. First, LLMs are primarily used to support decision-making and are not intended to be fully autonomous decision-makers. Most existing examples of the use of LLMs relate to tactical data-driven decisions involving scheduling, planning, maintenance, coordination, and other tasks in which LLMs assist in dealing with substantial amounts of data and making relevant decisions. At the same time, the growing role of human–LLM interaction is evident in many studies that use human-in-the-loop frameworks to integrate AI-generated results with human evaluation.

Despite these developments, several important tensions emerge in the literature. Although LLMs are frequently characterized as able to perform complex reasoning tasks, their low level of reliability and explainability reduce user trust and restrict their practical use. This creates a clear gap between the potential for autonomous decision-making and the ongoing need for human oversight. In addition, the use of general-purpose LLMs does not always align with the highly specific requirements of production environments, highlighting the need for better integration with domain knowledge and context-specific data. Furthermore, while many studies recognise the importance of human–AI interaction, only a few examine in detail how factors such as trust, usability, and cognitive load influence the effective use of LLMs in decision-making.

To conclude, the field still seems fragmented and at its early development stages, with most studies being conducted based on simulation or experimentation with no practical testing verification. Most research has concentrated on the technical aspects rather than the decision process and interaction mechanisms. This indicates that while there is great promise in using LLMs to facilitate production decisions, there is still much to learn about their functioning in practical decision-making scenarios.

6 Research Gap and Future Research Directions

Despite the growing interest in LLM-supported decision-making, several critical gaps remain in the literature. First, there is a lack of formalised models for human–LLM interaction. While human-in-the-loop methods are widely used, defining roles and interaction processes is difficult, hindering the development of decision support systems that can be designed, compared, and evaluated systematically (RQ1).

Secondly, the quality of decision-making and the performance of LLM-based decision-making are underexplored. Many studies focus on technical aspects, such as a machine's ability to provide decision support, but fail to present evidence of its contribution to improved decision outcomes. Consequently, essential characteristics, including accuracy and robustness, are not properly analysed, making it impossible to assess the true value of decision support provided by LLMs (RQ2).

Furthermore, there appears to be little research into domain-related adaptation and integration into production systems. A number of researchers employ LLMs that are general in nature without adequately considering the importance of domain knowledge, structured data integration, and interoperability with existing industrial systems. This limits the practicality of the current techniques in actual implementation scenarios.

In addition, the current body of literature remains relatively fragmented, with many studies focusing on specific applications or technical solutions rather than offering a comprehensive understanding of decision-making processes. A significant number of studies also lack real-world validation, which limits the ability to evaluate how LLM-supported decision-making performs in practical industrial environments. Moreover, the influence of human–LLM interaction and organisational factors on decision outcomes is still insufficiently explored.

Therefore, future studies should focus on ways to develop an appropriate structure for collaboration, define standardized approaches to assessing decision quality, and carry out empirical research on how these types of systems can be applied in industry settings. Greater emphasis is also needed on human-centred design and system integration, ensuring that LLM-based decision systems are not only technically effective but also usable, trustworthy, and aligned with organisational processes

7 Research Limitations

There are some limitations associated with the study. First, the review of literature was performed using scientific articles available only in English, indexed in Scopus and WoS databases. This means that relevant

studies published in different databases or conferences, or written in languages other than English, might be excluded from the analysis.

8 Conclusion

This study provides a structured literature review of emerging applications of large language models (LLMs) in decision-support systems in manufacturing contexts. The findings show that LLMs are widely used to support tactical decision-making tasks such as scheduling, planning, diagnostics, and coordination, offering clear advantages in data processing, efficiency, and reasoning capabilities. However, their actual influence on decision-making processes and outcomes remains insufficiently explored. Existing research tends to focus primarily on technological performance, with limited attention to how LLMs can be effectively integrated into decision-making processes, particularly to improve collaboration and acceptance of AI-generated outcomes.

The present study adds to existing knowledge by highlighting certain significant gaps. To start with, there is a lack of well-defined frameworks that explain how humans and LLMs should collaborate in decision-making contexts. Second, there is an absence of sufficient empirical information from actual production settings, thus constraining our understanding of how the system functions in practice. Third, there is a lack of studies examining the factors that influence human–LLM collaboration, including aspects such as trust, usability, and decision authority.

The results also indicate that the success of the integration of LLMs into decision support systems depends on the alignment between technological capabilities, decision-making requirements, and organisational processes. Without such alignment, the potential benefits of LLMs may not be fully realised.

Like any research, there are certain limitations to the current study. The review is based on academic articles published in English and indexed in the Scopus and Web of Science databases. As a result, relevant studies from other databases or non-English sources may not have been included, which may limit the completeness of the analysis and the generalisability of the findings.

Future research should therefore focus on developing structured frameworks for human–LLM collaboration, establishing clear criteria for evaluating decision quality, and conducting empirical studies in real industrial settings. These directions are essential for advancing the practical use of LLMs and ensuring their effective contribution to decision-making in manufacturing.

9 Ethic Declaration

The study is solely based on existing literature without engaging human participants or experimenting. Therefore, ethical approval is not necessary for the study.

10 AI Declaration

The tool Rayyan was used to assist in identifying and removing duplicate records during the study selection process. Additionally, Google Gemini was used to generate a PRISMA 2020 flow diagram illustrating the study selection process.

References

Ajraou, A. et al. (2026) “ACT-FLEX robotic decision architecture”, Robotics and Computer-Integrated Manufacturing.

Cai, Y. et al. (2025) “A decision-making framework for human–LLM collaboration in smart grids”, Preprint.

Cariaga, J. (2025) “Human-AI collaboration in decision-making, creativity and productivity: systematic review”, Journal of Ongoing Artificial Intelligence Innovations.

Cimino, A. et al. (2026) “Leveraging large language models to enhance simulation-based decision-making in manufacturing systems”, European Journal of Innovation Management.

Colabianchi, S. et al. (2024) “Assessment of a large language model based digital intelligent assistant in assembly manufacturing”, Computers in Industry.

Ding, L. et al. (2026) “LLM-powered human-robot collaboration agent”, Robotics and Computer-Integrated Manufacturing.

Dong, X. et al. (2025) “Physical-to-virtual human-robot interaction”, Manufacturing Letters.

Elgeddawy, T. (2026) “Human–AI collaboration models in supply chain decision-making”, International Journal of Scientific Research in Science, Engineering and Technology.

Fan, Z. et al. (2024) “Embodied intelligence in manufacturing: leveraging large language models for autonomous industrial robotics”, Journal of Intelligent Manufacturing.

Fan, Z. et al. (2026a) “A large language model-based decision support framework for smart manufacturing execution systems”, Computers and Industrial Engineering.

Fan, Z. et al. (2026b) “LLM-guided reinforcement learning framework”, Robotics and Computer-Integrated Manufacturing.

Favier, M. et al. (2025) “Leveraging LLMs for collaborative human-AI decision making”, AAAI Spring Symposium Series.
Garcia, R. et al. (2024) “Framework for LLM applications in manufacturing”, Manufacturing Letters.
Genetti, E. et al. (2026) “Enhancing decision-making in production scheduling through digital twin and explainable AI integration”, International Journal of Production Economics.
Gholami, S. (2024) “Artificial intelligence for sustainable reconfigurable manufacturing systems: a decision support framework”, Big Data and Cognitive Computing.
Huang, Y. et al. (2026) “LLM-enhanced knowledge service for process planning”, Journal of Manufacturing Systems.
Ioshchikhes, M. et al. (2026) “Knowledge engineering for manufacturing decision support using large language models and fuzzy logic systems”, AHFE.
Jadhav, P. et al. (2025) “LLM-3D print: large language models to monitor and control 3D printing”, Additive Manufacturing.
Kitchenham, B. and Charters, S. (2007) Guidelines for performing systematic literature reviews in software engineering, EBSE Technical Report, Keele University, UK.
Li, X. et al. (2025) “Large manufacturing decision model (LMDM)”, CIRP.
Li, X. et al. (2026) “AI-agent system with MKG-RAG for 3D printing”, Robotics and Computer-Integrated Manufacturing.
Lin, H. et al. (2025) “5D digital twin with LLM-enhanced reinforcement learning”, Robotics and Computer-Integrated Manufacturing.
Liu, H. et al. (2026a) “LLM-enhanced embodied multi-agent manufacturing system”, Journal of Manufacturing Systems.
Liu, H. et al. (2026b) “Integrating human intelligence and LLMs in AR-assisted motion planning”, Robotics and Computer-Integrated Manufacturing.
Marandi, A. et al. (2025) “A large language model and knowledge graph-based framework for fault diagnosis in industrial systems”, Applied Sciences.
Naqvi, M. et al. (2025) “Semantic similarity-based industrial maintenance retrieval using BERT and ontology integration”, Journal of Industrial Information Integration.
Nie, S. et al. (2026) “Industrial knowledge-enhanced fault diagnosis method integrating LLM and knowledge graph”, Computers and Industrial Engineering.
Ouzzani, M., Hammady, H. and Fedorowicz, Z. et al. (2016) “Rayyan—A web and mobile app for systematic reviews”, Systematic Reviews, Vol. 5, No. 1, pp 210.
Page, M.J., McKenzie, J.E. and Bossuyt, P.M. et al. (2021) “The PRISMA 2020 statement: an updated guideline for reporting systematic reviews”, BMJ, Vol. 372, n71.
Passerini, K. et al. (2025) “Fostering effective hybrid human-LLM reasoning and decision making”, Frontiers in Artificial Intelligence.
Rank, Y.D. et al. (2025) “Leveraging large language models for tacit knowledge capture and transfer in manufacturing systems”, AHFE.
Sabbah, R. and Li, Y. (2025) “When humans and large language models collaborate, problem-finding illuminates”, Innovation: Organization and Management.
Sadeghi, A. et al. (2026) “The AI-driven decision-making framework integrating AHP and ChatGPT for supplier selection”, Annals of Operations Research.
Skrzek, M. et al. (2026) “Ontology-based knowledge integration for complex product manufacturing using large language models”, IFIP.
Sridhar, V. et al. (2026) “A comprehensive framework for human-AI collaborative decision making”, Expert Systems with Applications.
Tan, J. et al. (2026) “Adoption of large language models in supply chain and production planning”, International Journal of Production Economics.
Teng, Y. et al. (2025) “Cerebrum twin (6D digital twin)”, Journal of Manufacturing Systems.
Tong, Z. et al. (2026) “LLM-enabled carbon emissions abatement in green production”, International Journal of Production Economics.
Vafeiadis, T., Artiba, A. and Urbanic, M. (2025) “Large language models in manufacturing: a survey of applications, capabilities and challenges”, IEEE Conference.
Werheid, M. et al. (2025) “Designing an LLM-based copilot for manufacturing equipment selection”, Manufacturing Letters.
Wu, D. et al. (2024) “ProcessCarbonAgent: a large language models-empowered autonomous agent”, Journal of Manufacturing Systems.
Wu, D. et al. (2025a) “Multi-agent reinforcement learning-based scheduling optimization enhanced by LLMs”, Computers and Industrial Engineering.
Wu, D. et al. (2025b) “Fault identification with LLM and knowledge graph”, International Journal of Advanced Manufacturing Technology.
Wu, D. et al. (2026) “Causality-driven collaborative optimization with LLM-powered multi-agent systems”, Journal of Manufacturing Systems.
Wulf, J. et al. (2024) “Understanding the value and risks of large language models in manufacturing organizations”, Springer.
Yang, F. et al. (2026a) “Multi-robot planning and coordination”, Robotics and Computer-Integrated Manufacturing.
Yang, F. et al. (2026b) “Decentralised task planning and motion coordination for scalable manufacturing”, Robotics and Computer-Integrated Manufacturing.
Yan, X. et al. (2026) “Knowledge graph and LLM integration for process base construction”, Journal of Manufacturing Systems.

Co-Creation of Knowledge in Marketing: Human-AI Collaboration in the Construction of Proto-Personas

Eduardo Trauer
State University of Santa Catarina – ESAG/Udesc, Florianópolis, Brazil

Abstract

This study investigates, through the lens of the Knowledge Management 5.0 (KM 5.0) paradigm, the role of Generative Artificial Intelligence (GAI) as a partner in the co-creation of knowledge for the development of proto-personas within the context of data-driven marketing. The research addresses the challenge of understanding consumer behaviour in niche markets – in this study, the bonsai market – which are characterised by a scarcity of structured data and a predominance of tacit context.

An applied qualitative approach was adopted, with an exploratory-descriptive design. The study was conducted with 71 undergraduate students (67 valid responses; response rate of 94.4%) enrolled in Marketing courses at ESAG/Udesc. Participants interacted with a Generative Artificial Intelligence Agent (GAIA) named John Jobs Naka (J. J. Naka), configured using a specialised corpus derived from in-depth qualitative interviews with recognised sector experts, together with technical publications and secondary data. The methodological framework was structured on the SECI model (Socialisation, Externalisation, Combination and Internalisation) to enable the analysis of knowledge conversion processes in human-AI interaction.

The findings demonstrate that GAIA can function as a cognitive mediator, facilitating the externalisation and integration of knowledge while supporting the development of proto-personas as strategic marketing artefacts. The quantitative results indicate high levels of acceptance and perceived utility – 98.5% of participants recognised the agent as an effective partner; the clarity score was 4.19/5; and recommendability 8.57/10. Qualitative analysis highlights three key aspects: cognitive mediation, empathy and humanisation (aligned with H2H marketing), and epistemological integration of tacit and explicit knowledge.

This cutting-edge study represents an advancement of KM 5.0 by demonstrating the capacity of GAIA to enhance learning processes, support agile knowledge development and enable the generation of sophisticated consumer insights in niche markets. The findings also have relevant implications for the fields of marketing and higher education, positioning GAIA as scalable tools for the co-creation of knowledge while preserving human centrality in decision-making processes.

Keywords: Generative Artificial Intelligence Agent; Knowledge Co-creation; Proto-personas; Data-driven marketing; KM 5.0.

Paper type: Academic Research Paper

1 Introduction

The development of proto-personas in the context of data-driven marketing has emerged as a relevant exploratory practice for understanding consumers and markets, particularly in niche contexts characterised by a high degree of specialisation, tacit knowledge, and a lack of structured data. The bonsai segment is a typical example of such a niche.

In educational and organisational settings, efficient consolidation of quantitative and qualitative information, frequently drawn from specialised and difficult-to-access sources, is essential. GAI, when informed by a refined knowledge base, may assume the role of a supporting agent for the co-creation of market knowledge. In this study, the findings from in-depth qualitative interviews with recognised bonsai experts constituted the knowledge base. The generation of proto-personas by undergraduate Business Administration students, using a GAIA, illustrates a dynamic of human AI collaboration. Thus, interpretation, synthesis and contextualisation are not entirely delegated to technology, but rather constructed jointly, combining human expertise, empirical data and generative capacity of the system.

The aim of this article is to analyse how collaboration between ESAG/Udesc students and GAIA – configures on the basis of specialised knowledge co-creation in marketing through the construction of proto-personas that represent the ideal customer profile in the target market. The research is guided by the following question: how can human-AI collaboration mediated by GAIA grounded in specialised knowledge support the agile generation of proto-personas within the context of data-driven marketing?

2 Literature Review

The use of buyer personas and proto-personas in marketing has been discussed as a practice aimed at understanding consumers and achieving strategic alignment of teams, products and markets. While buyer personas require consolidated databases, proto-personas emerge as preliminary representations constructed from the empirical knowledge of specialist – i.e., the externalisation of their tacit knowledge (Cooper, 2004; Nonaka and Toyama, 2021; Pinheiro et al., 2018; Trauer and Valdati, 2023; Trauer and Valdati, 2024; Trauer and Valdati, 2025). This approach proves particularly relevant in niche markets where structured data are scarce, and knowledge resides predominantly in the accumulated experience of actors in the sector. In these contexts , the construction of proto-personas assumes an exploratory and iterative character, enabling the externalisation of implicit knowledge and the initial formulation of patterns of consumer behaviour, motivations and needs, as reported in previous research on knowledge-creation and use in organisations (Davenport and Prusak, 1998).

Advances in the field of GAI and Large Language Models (LLMs) have broadened the debate on the role of these technologies as mediators in complex cognitive processes. Unlike traditional decision-support systems, GAIAs are capable of synthesising information from multiple sources, structuring narratives and supporting interpretive processes in direct interaction with human users. Indeed, AI agents and generative AI have been shown to be at the peak of the hype cycle, exhibiting strong momentum in the current market (den Hamer et al., 2025; Gartner, 2024; Mora-López et al., 2026). These systems may function as cognitive partners, particularly when supplied with specialised and contextualised knowledge bases (Trauer and Valdati, 2024; Trauer and Valdati, 2025). In the field of marketing, this human-AI collaboration has been associated with the expansion of analytical capacity, the acceleration of exploratory processes, and the strengthening of empathy in the understanding of specialist markets, reinforcing human-centred relational approaches as presented by H2H paradigm (Kotler et al., 2024).

In the context of education and data-driven marketing, agile methodologies have gained prominence as particularly well-suited to addressing problems, characterised by uncertainty, complexity, and constraints of time and resources. The literature suggests that short cycles of experimentation, incremental learning and continuous validation are especially effective in the construction of exploratory artefacts such as proto-personas (Cooper, 2004; Trauer and Valdati, 2024). The integration of agile methodologies with GAIA enhances this process by enabling the rapid incorporation of specialised knowledge, the interative reformulation of hypotheses, and the reduction of costs associated with traditional qualitative data collection. Asa such, collaboration between humans and GAIA, underpinned by knowledge bases derived from human experts, constitutes a promising approach to the co-creation of market knowledge in niche markets, thereby contributing to consistent, efficient and contextualised decision-making.

The transition from conventional GAI to agent-based architecture represents a fundamental evolution, moving from systems that merely respond to commands towards proactive computational entities. The following section develops this theoretical framework with a focus on the integration of RAG-type (Retrieval-Augmented Generation) agents.

2.1 The nature of generative AI agents

GAI differs from traditional algorithms in that operates in a probabilistic and non-deterministic manner. Whereas an algorithm follows rigid coding steps to transform an input into an output, GAI employs neural networks trained on vast datasets to identify patterns and generate new content(Dar and Gul, 2025).

When this generative capacity is encapsulated in the form of an AI agent, the system acquires autonomy. Historically, AI research has focused on the development of rational agents that act to achieve the best expected outcome in the face of uncertainty (Russell et al., 2021). Currently, an AI agent is defined as an autonomous or semi-autonomous software entity that employs AI techniques to perceive its environment, make decisions, execute actions and attain goals.

These authors differentiate the terminology along a spectrum of autonomy as follows:

- AI assistants: modules that assist with tasks under the direct request and guidance of a human user;
- AI agents: entities that operate independently, extending beyond simple responses to prompts by assuming control of planning and execution to achieve objectives; and
- Agentic AI: a broader approach to building solutions based on one or more collaborative AI agents.

2.1.1 Core architecture and agentic capabilities

Unlike a standalone LLM, and agent combines the LLM with peripheral components in order to act (Taurion, 2026). The anatomy of an AI agent is structured around six core capabilities: perception, decision-making, actioning, agency, adaptability and knowledge (den Hamer et al., 2025).

LLMs function as the central cognitive engine for interpretation and decision-making, and their architecture comprises:

- Memory: essential for maintaining the context of ongoing tasks and retaining the history of interactions (Taurion, 2026);
- Orchestration layer (planning): controls the execution flow, frequently operating in continuous loops (primary execution loops, secondary adaptation loops, and tertiary control/planning loops) (den Hamer et al., 2025); and
- Tools and connectors: interfaces that enable the agent to interact with the real world, query APIs, analyse documents and retrieve information from external databases (Taurion, 2026).

2.1.2 RAG-Based agents (Retrieval-Augmented Generation)

Among the greatest challenges for the adoption of LLMs in corporate and business environments are the phenomenon of hallucination (the fabrication of data) (Kalai et al., 2025; OpenAi, 2025; Swain, 2025; Xing, 2025), and the lack of transparency (the so-called *black box* problem) in the decision-making process of models. To mitigate these risks and ensure that the agent's knowledge is factual, up to date and domain-specific risks and ensure that the agent's knowledge is factual, up to date and domain-specific, RAG-type agent architecture may be employed (Acatrinei, 2025; Wang et al., 2026).

RAG architecture functions by anchoring the agent's responses in domain-specific knowledge bases, rather than relying solely on the model's parametric (trained) knowledge, thereby mitigating hallucinations and improving reliability in decision-making workflows (Wang et al., 2026). According to Wang et al. (2026), the structure of a knowledge management focused agent employing RAG functions as presented in Figure 1:

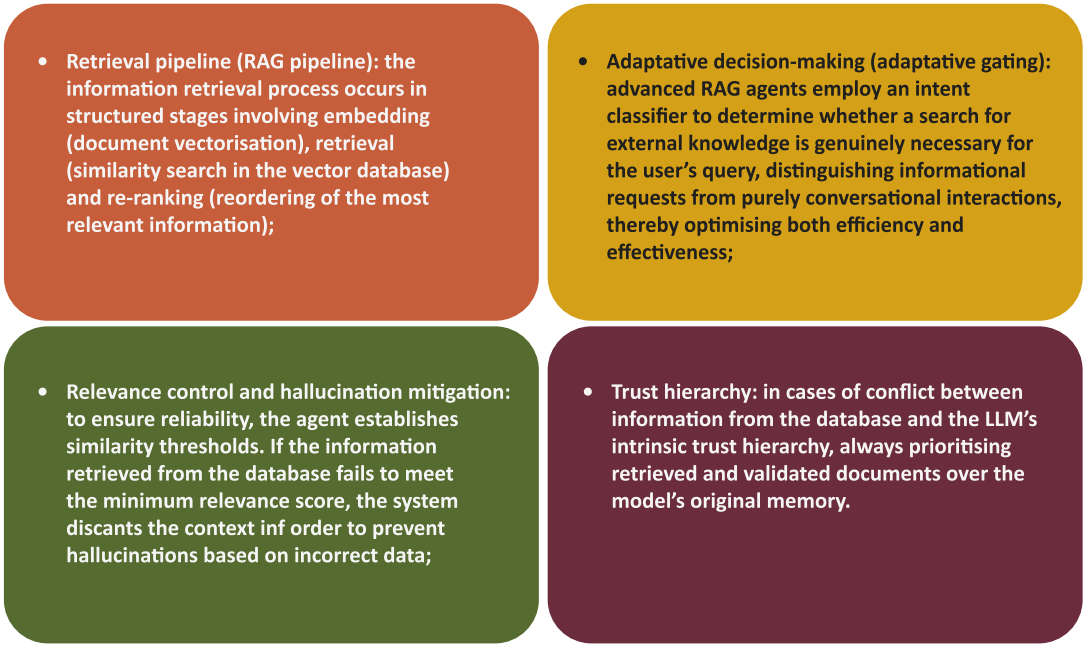


Figure 1- RAG-Based agents (Retrieval-Augmented Generation) (Wang et al., 2026)

In the context of marketing and proto-persona ideation, the RAG agent does not operate by generating creative text from nothing but instead retrieves factual data on consumer behaviour, guidelines provided, and reports from the database supplied to it, and subsequently employs the LLM to analyse and synthesise this information. This ensures that human-AI co-creation is simultaneously innovative and grounded in the empirical reality of the market, drawing on the interviews conducted and provided to GAIA (Acatrinei, 2025; Wang et al., 2026).

2.2 Human-AI collaboration in the construction of proto-personas

Proto-personas are semi-fictional representations of ideal customers, conceived on the basis of tacit knowledge and the insights shared by specialists (Trauer and Valdati, 2023; Trauer and Valdati, 2024; Trauer and Valdati, 2025). They offer an agile ana low-cost approach to strategic alignment, hypothesis formulation, and

consumer-centred decision-making, enabling rapid actions while more robust market data are collected and validated.

The incorporation of GAIA into this process acts as a cognitive catalyst. Rather than functioning as a passive tool, the AI operates as a cognitive mediator and empathic interlocutor, facilitating the externalisation of practitioners' knowledge and its internalisation by undergraduate students. The successful construction of these proto-personas occurs through the advantageous integration of the stages of the knowledge creation model proposed by (Nonaka and Takeuchi, 1995; Nonaka and Toyama, 2021): Externalisation – produced during the researcher's interviews with sector specialists; Internalisation – during the students interviews process with J. J. Naka; Socialisation once again – arising from reports produced by students. Socialisation also occurs when students present their findings orally in seminars.

This collaborative dynamic reflects the principles of H2H marketing (Kotler et al., 2024), ensuring that, despite algorithmic mediation, the process retains authenticity and remains centred on the real and psychological needs of individuals. The interative interaction with AI provides a safe environment, available around the clock, for testing, simulating real market conditions, and deepening target audience segmentation.

2.3 Challenges, ethics and corporate governance

Agent-mediated co-creation shows considerable promise. However, it presents significant practical and ethical challenges. The increasing autonomy of AI calls for careful attention to reliability and security (Taurion, 2026). Language models may make interpretive errors of fabricate plausible yet incorrect facts, giving rise to a phenomenon known as hallucination (Dar and Gul, 2025; Kalai et al., 2025; OpenAi, 2025; Swain, 2025; Taurion, 2026; Xing, 2025). Furthermore, because of their inherent complexity, deep learning algorithms frequently operate as a black box and therefore undermine transparency in decision-making (Dar and Gul, 2025).

Excessive delegation of critical thinking and creativity to AI may lead to dependence on technology, compromising the problem-solving abilities of the humans involved. The security of corporate data is equally critical, as agents that interact with and retrieve information from real databases and APIs are vulnerable to attacks such as *prompt injection*, which may force them to leak sensitive data or make unintended decisions (Dar and Gul, 2025; Taurion, 2026).

Implementing these corporate architectures requires more than mere technological adoption – it requires rigorous governance. It is essential to establish clear boundaries for AI operation, to maintain human supervision and validation at critical stages, to ensure continuous monitoring and to enforce stringent data protection policies. With appropriate regulation and oversight, GAIAs may serve as secure and scalable partners in the generation of insights and market intelligence.

3 Methodology

This study adopts an applied qualitative approach with an exploratory-descriptive design, grounded in the KM 5.0 paradigm. This emphasises the integration of advanced technologies with human-centred knowledge processes (Keidanren, 2018; Nonaka and Takeuchi, 1995; Nonaka and Toyama, 2021). The research was conducted within the Digital Marketing Environments and Data-Driven Marketing modules at ESAG/Udesc during the 2025-2 semester, incorporating a GAIA developed via ChatGPT. The agent named J. J. Naka, was designed to function as a cognitive mediator in knowledge co-creation processes, supporting the development of proto-personas in an educational context.

The target population that has participated in the final evaluation survey comprised 71 students enrolled in the 2025-2 semester, of whom 67 participated effectively, yielding a response rate of 94.4%. Participants included both Brazilian students and European exchange students, representing a heterogenous learning environment. Students conducted individual interviews with GAIA in either Portuguese and English and recorded the responses for subsequent systematisation. Employing a methodological framework previously introduced in the classroom, students organised themselves into groups to develop proto-personas representative of consumer segments within the Brazilian bonsai market. These proto-personas were constructed through the integration of tacit knowledge from real market specialists – externalised to form J. J. Naka knowledge base – together with secondary data, following the conceptualisation proposed by (Trauer and Valdati, 2024). Variables such as demographic profile, motivations, objectives, challenges and information sources were considered.

The methodological approach was based on the SECI model of knowledge creation – Socialisation, Externalisation, Combination and Internalisation – originally proposed by (Nonaka and Takeuchi, 1995).

Data collection was carried out by means of a structured questionnaire comprising both closed-ended items (Likert scale) and open-ended questions. This enabled the assessment of variables such as the clarity of the agent's responses, its contribution to understanding consumer behaviour, the perception of the agent as a

cognitive partner, and the intention of use AI in the future. The findings indicate that integration of generative AI with active pedagogical strategies enhances the development of proto-personas as both a learning tool and strategic marketing resource in niche markets. Furthermore, the results reinforce the relevance of KM 5.0 principles and H2H marketing (Kotler et al., 2024), emphasising the rote of AI as a facilitator in the co-creation of knowledge without compromising human centrality in the learning process.

4 Results

The response adhesion rate in this survey was 94.4%, derived from 67 valid responses. This figure approaches census-level coverage. Amongst the principal findings, the following merit particular attention:

- Quantitative efficacy and technological acceptance: data measured by Likert scale attest to a high level of acceptance and perceived pedagogical utility of the technology. Notably, 98.5% of participants recognised J. J. Naka as an effective mediator, perceiving the system not as a mere search engine but as an active partner in knowledge construction and a facilitator for the generation of detailed consumer insights. In addition, the agent's communicative capacity received a score of 4.19 out of 5 on the response clarity criterion. Satisfaction levels culminated in a strong behavioural intention, with the agent achieving a recommendability score of 8.57 out of 10.
- Qualitative foundations of successful co-creation: analysis of the open-ended questions and students' perceptions revealed that the efficacy of proto-persona construction rested upon three analytical pillars:
    - Cognitive mediation: J. J. Naka as a cognitive catalyst, facilitating the process of knowledge externalisation by students. This mediation promoted higher level strategic reflection on consumer behaviour within the Brazilian bonsai niche market, and resulted in more sophisticated and segmented market analyses;
    - Empathy and humanisation (H2H marketing): corroborating the theoretical framework of human-centred relational marketing, participants attributed markedly anthropomorphic characteristics to the system. The agent was described as practical and direct, instructive and attentive as well as empathic and trustworthy. This personification of technology proved fundamental in establishing trust and enabling authentic dialogue, both essential to the co-creation dynamic;
    - Epistemological integration: the collaboration enabled the advantageous integration of a technical and informational knowledge (supplied and structured by the AI) with tacit knowledge (inferred and produced by the student). This ensured the generation of proto-personas that were both market coherent and contextually aligned.
- Educational impact and practical skills development: in the context of the teaching and learning process, AI mediation strengthened active methodologies by creating a safe readily available environment for experimentation. Students were able to iterate, test hypotheses and refine their interpretations of proto-personas through continuous feedback cycles. Consequently, the development of practical skills for transforming abstract theories into actionable insights was observed simulation real market conditions and enhancing data-driven analytical capabilities. The motivational impact was also considerable, with multiple students reporting their intention to incorporate such agents as strategic allies both in their professional lives and in future projects.

These experiment findings reinforce the premise that GAIAs can function effectively as empathic interlocutors and co-creators of value, enabling pedagogical innovation and knowledge management practices in specialised niche markets.

4.1 Contributions to academia

In the academic and educational context, the construction of proto-personas supported by GAIA represents a significant technological and pedagogical innovation. This integration contributes to learning in several noteworthy ways:

- Fostering active methodologies and safe experimentation: GAIA-mediated interaction provides a safe environment for experimentation where students may test hypotheses about consumer profiles and iteratively refine their interpretations on a continuous basis;
- Development of practical skills and critical thinking: the use of these technologies helps students to transform abstract theory into actionable insights. By simulating real market conditions in the creation of proto-personas, students develop data-driven analytical capabilities and sharpen their critical thinking skills;

- Experiential and applied learning: for higher educational institutions, the systematic adoption of AI agents as pedagogical mediators in field work activities, interdisciplinary projects or market simulations directly promotes high-impact experiential applied learning;
- Cognitive partnership in knowledge co-creation: AI agents function as cognitive partners and empathic interlocutors, actively assisting in the construction and structuring of the knowledge required to design proto-personas;
- Increased motivation and professional development: the use of AI in the classroom tends to significantly increase student engagement and motivation in future projects. The practice actively prepares future professionals, particularly in marketing, by cultivating a culture of co-creation and innovation while maintaining human centrality and autonomy in the learning process.

4.2 Contributions to the market

Beyond the educational environment, proto-personas are employed as provisional, semi-fictional representations of ideal customers. They are constructed from the tacit knowledge and shared insights of specialist teams. Originally conceived in the software industry as a low-cost, rapid-implementation alternative (Pinheiro et al., 2018), they are being adapted to the marketing environment as a means of generating tactical and strategic marketing actions (Trauer and Valdati, 2023).

Their practical application in the corporate and business world is most prominent in the following areas:

- Strategic alignment and decision-making: proto-personas provide an agile approach to aligning teams, jointly formulating hypotheses about customers and guiding strategic decisions. They offer an immediate, consumer-centred perspective whilst more robust and detailed market data are collected for future validation;
- Competitive advantage in marketing: for marketing professionals, the construction of proto-personas, particularly when optimised through GAIA, serves as tactical tool for cost reduction and increased strategic response speed. They are especially useful in the initial exploration of as yet unstructured market segments and consumption niches that demand a high degree of customisation, and
- Multidisciplinary adoption: the logic of proto-personas of the detailed understanding of users' needs may be extended beyond technology and marketing, to sectors such as design, business administration, humane-centred smart cities, and healthcare.

Figure 2 presents an adaptation of the SECI model of knowledge creation proposed by Nonaka and Takeuchi (1995) and Nonaka and Toyama (2021), applied to the context of proto-persona development supported by GAIA – here referred to as J. J. Naka. The model retains the classical structure of the four knowledge conversion modes – Socialisation, Externalisation, Combination and Internalisation. It incorporates empirical elements from this research process, highlighting technological mediation as an integral component of the knowledge creation dynamic. At the centre of Figure 2, agent Naka functions as an articulating element, connecting the different phases of the process and enhancing the transformation and circulation of knowledge amongst participants.

In the Socialisation quadrant, the origin of tacit knowledge is identified through in-depth interviews with bonsai market specialists, as well as the continuation of this process in student interactions and group discussions. The inclusion of seminar presentations within this quadrant reinforces the notion that the exchange of experiences, perceptions and interpretations between students and the lecturer constitutes a moment of shared tacit knowledge generation.

In the Externalisation quadrant, this knowledge becomes explicit through the recording and transcription of interviews, and the incorporation of secondary sources such as books and technical reports that feed into the GAIA's knowledge base. This process illustrates the structured transformation of tacit knowledge into communicable and reusable artefacts.

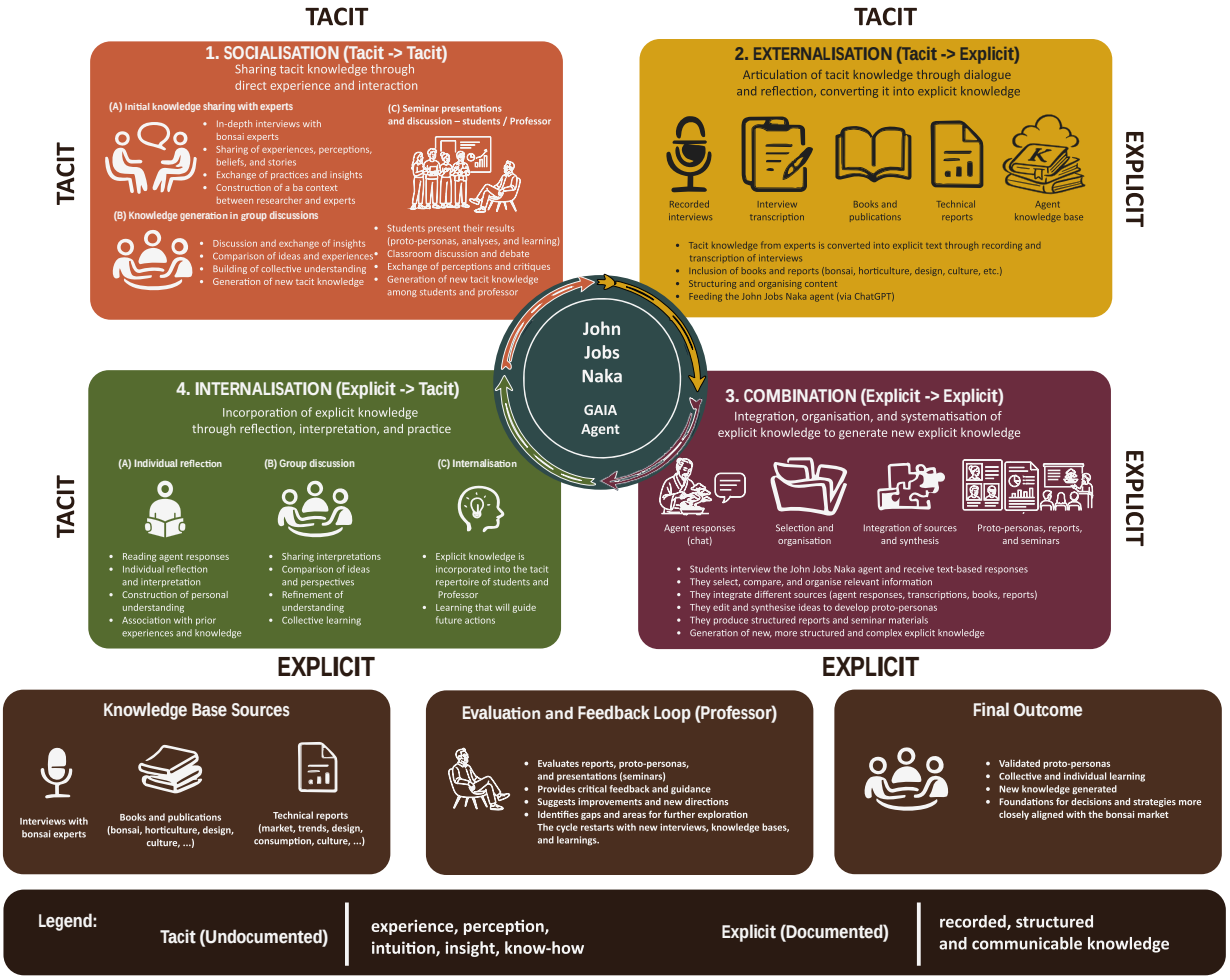


Figure 2 - Knowledge Creation Model - Proto-Personas with GAIA adapted by Nonaka and Toyama (2021)

In the lower quadrants, Figure 2 makes explicit an analytical distinction between Combination and Internalisation. Combination is represented by the activities of organising, selecting, integrating and synthesising explicit content. These are carried out by students on the basis of the agent's responses and other sources, and generate proto-personas, reports and presentation materials. Internalisation occurs when students individually read, reflect upon and interpret this information, incorporating it into their tacit repertoire - a process that extends to group discussions and the learning arising from participating in seminars. Finally, the assessment conducted by the lecturer serves as a feedback mechanism for the system, which contributes to the continuous refinement of knowledge and reinforces the spiral nature of the model, consistent with the dynamic perspective of organisational knowledge creation.

4 Limitations

The configuration of the John Jobs Naka agent via ChatGPT presented specific structural characteristics and limitations during its application.

- Closed knowledge base: the agent was configured with a restricted corpus (knowledge base), which may limit the exploration of content extending beyond the specific scope of bonsai;
- Absence of cross-session memory: the system employed does not retain memory of interactions across different sessions, which restricts the capacity for medium- and long-term longitudinal knowledge construction;
- Lack of external validation: the project did not include external validation by bonsai segment specialists to confirm the accuracy and level of expertise of the domain information; and
- Refinement limitations: time constraints during the academic semester means there were limited opportunities to refine configuration prompts, test new iterations or incorporate student feedback with a view to improving the agent's responses.

5 Conclusion

This study demonstrates that human-AI collaboration mediated by GAIA and underpinned by specialist expertise may function as an efficient and effective mechanism of co-creation within a data-driven marketing context. It may enable the agile generation of proto-personas in niche markets characterised by a scarcity of structured and secondary data. From the perspective of the SECI model, the findings indicate that the GAI agent may operate as a cognitive mediator, enhancing the processes of externalisation, combination, and internalisation of knowledge, without displacing the centrality of human interpretation.

Empirically, the interaction between students and the agent J. J. Naka facilitated the development of coherent and contextually-grounded artefacts without the need for physical travel or scheduling consultations with human experts. The process enabled the integration of specialised tacit knowledge and explicit information through continuous cycles of refinement. This dynamic supports the notion that GAI agents, when anchored in refined knowledge bases, can expand analytical capacity and reduce divergences in the exploration of niche markets.

As a contribution, the study integrates KM 5.0 with agent-based generative AI architectures, highlighting their applicability in the construction of marketing knowledge - particularly proto-personas - and in supporting active learning methodologies. From a methodological standpoint, it demonstrates the feasibility of employing a GAIA as a structuring instrument for the generation of exploratory artefacts. In practical terms, the findings point to the potential for accelerating analytical and decision-making processes in contexts characterised by high complexity and limited data availability.

The limitations identified indicate the need for external validation, expansion of the knowledge corpus, and application across different contexts, thereby opening avenues for future research into the robustness and scalability of human-AI co-creation.

6 Ethic declaration

The author declares that the research described in this study involved human participants in an educational context, characterising it as a low-risk activity. Student participation was voluntary, and the data collected were treated anonymously and confidentially, with no identification of participants at any stage.

In accordance with institutional guidelines and the nature of the study, submission to a Research Ethics Committee was not required. The study upheld the ethical principles applicable to research involving human participants, including confidentiality, voluntariness and the responsible use of data.

7 AI declaration

The author declares that GAI tools were employed in a limited and controlled manner during the development of this study. Specifically, AI was used to support the initial development of Figure 1, inspired by the SECI model proposed by Nonaka and Takeuchi (1995) and Nonaka and Toyama (2021).

The AI tool was employed in the visual structuring and preliminary layout of this figure, including suggestions regarding spatial organisation, graphic elements and visual hierarchy. The material generated was used exclusively as a draft for conceptual visualisation. Figure 1 was subsequently refined by the author using Adobe Illustrator software.

All subsequent stages - including the refinement of the model, theoretical alignment, adaptation to the empirical context, and the final version of Figure 1 - were developed, critically evaluated and validated in their entirety by the author, thereby reflecting the author's intellectual contribution and theoretical interpretation.

References

Acatrinei C (2025) To Use or Not to Use Artificial Intelligence (AI) in Marketing: Insights from Qualitative Research on Adopters and Non-adopters. *Contributions to Management Science*. Springer Science and Business Media Deutschland GmbH, pp.253–267.

Cooper A (2004) *The Inmates Are Running the Asylum: Why High-Tech Products Drive Us Crazy and How to Restore the Sanity.* Indianapolis, Indiana: Sams Publishing.

Dar AR and Gul M (2025) Generative AI in Consumer Behaviour and Marketing: Applications, Opportunities and Challenges. In: Skeja A, Ramadani V, Sadiku-Dushi N, et al. (eds) *Navigating AI in Business: Strategies and Insights Across Disciplines*. Cham: Springer Nature Switzerland, pp.259–274.

Davenport TH and Prusak L (1998) *Working knowledge: How organizations manage what they know.* Harvard Business School Press.

den Hamer P, Ramos L and Brethenoux E (2025) When to Use or Not to Use AI Agents. Reportno. Report Number|, Date. Place Published|: Institution|.

Gartner I (2024) See 3 of the Most Popular Gartner Hype Cycles. Reportno. Report Number|, Date. Place Published|: Institution|.

Kalai AT, Nachum O, Vempala SS, et al. (2025) Why Language Models Hallucinate. *arXiv preprint*.

Keidanren (2018) Society 5.0: Co-creating the future. Reportno. Report Number|, Date. Place Published|: Institution|.

Kotler P, Pfoertsch W, Sponholz U, et al. (2024) *Marketing H2H: a jornada para o marketing human to human.* São Paulo: Benvirá.

Mora-López JP, Lopez-Lopez D and Rivera-Hernaez O (2026) Unveiling the Generative AI boom: what hype metrics reveal for digital business and E-commerce. *Electronic Commerce Research* 26(2): 2423–2451.

Nonaka I and Takeuchi H (1995) *The knowledge-creating company: How Japanese companies create the dynamics of innovation.* Oxford University Press.

Nonaka I and Toyama R (2021) The knowledge-creating theory revisited: Knowledge creation as a synthesizing process. In: Kessler EH (ed) *Encyclopedia of management theory*. Sage, pp.421–426.

OpenAi (2025) Why language models hallucinate. *OpenAI*.

Pinheiro EG, Lopes LA, Conte TU, et al. (2018) The contribution of non-technical stakeholders on the specification of UX requirements: An experimental study using the proto-persona technique. In: Gottardi T (ed) *32nd Brazilian Symposium on Software Engineering, SBES 2018.* Association for Computing Machinery, 92–101.

Russell S, Norvig P, Chang M-W, et al. (2021) *Artificial intelligence : a modern approach.* Hoboken, N.J: Pearson.

Swain G (2025) OpenAI admits AI hallucinations are mathematically inevitable, not just engineering flaws. *Computerworld*.

Taurion C (2026) Em um recente evento, percebi que muitos falavam em agentes de IA, mas tinham dificuldade em explicar exatamente o que são. *LinkedIn*.

Trauer E and Valdati AdB (2023) PROTO-PERSONAS: AGILIDADE NA IDENTIFICAÇÃO DO CLIENTE IDEAL. *Anais do Congresso Internacional de Conhecimento e Inovação – ciki* 1(1).

Trauer E and Valdati AdB (2024) Proto-Personas: Innovation and Agility in the Process of Identifying the Ideal Customer. In: Morales-Alonso G and Schiuma G (eds) *19th International Forum on Knowledge Asset Dynamics (IFKAD).* Madrid, Spain: Institute of Knowledge Asset Management (IFKAD), 225–241.

Trauer E and Valdati AdB (2025) Generative AI as a Knowledge Co-Creation Partner: Building Proto-Personas in Data-Driven Marketing under the KM 5.0 Paradigm. In: *GFIC 2025*, Lisbon, Portugal. ICAA.

Wang Q, Ma H, Hu J, et al. (2026) Multi-agent cooperation for smart gas reservoir management. *Expert Systems With Applications* 319.

Xing W (2025) Why OpenAI's solution to AI hallucinations would kill ChatGPT tomorrow. *The Conversation*.

Beyond Prediction: Building Sustainable Human-AI Knowledge Ecosystems in Healthcare Organizations

Sara El Matouk, Daniele Allega
Universitas Mercatorum, Rome, Italy

Abstract

Healthcare organizations increasingly rely on unstructured narrative material - such as clinical notes, referral letters, discharge summaries, and other health-related texts - to understand patient trajectories, treatment responses, and emerging vulnerabilities. Yet the sustainable use of these resources cannot be reduced to predictive performance. This paper argues that the long-term viability of AI in healthcare depends on the socio-technical and governance conditions under which narrative data, model outputs, and human oversight are organized. It addresses the following question: what socio-technical and governance conditions are required for the sustainable development of human-in-the-loop AI systems based on health-related narrative data?

The argument is developed through PRAXIS, an early-stage doctoral research protocol that should be understood here as an illustrative governance case rather than as a validated clinical deployment. PRAXIS is designed to support signal extraction and interpretive structuring from Italian health-related narratives while keeping annotation and classification decisions under expert human supervision. In its current phase, the protocol relies on a corpus of participant-generated narratives from university students aged 19-24 and does not report final empirical results or clinician-facing deployment. This bounded setting is methodologically useful because it allows annotation procedures, interpretive standards, and governance mechanisms to be developed before extension to more sensitive populations and institutional settings.

The paper proposes a governance-first framework organized around three interdependent layers: an AI layer responsible for signal extraction and candidate annotation generation; a human layer responsible for review, disagreement management, calibrated reliance, and interpretive accountability; and an organizational knowledge governance layer responsible for provenance, access, lifecycle management, and controlled transfer across populations. Within this framework, label fitting - the embedding of population-specific assumptions in annotation schemes - is introduced as a distinct governance risk that differs from technical distribution shift and therefore requires lifecycle governance and cross-population review.

The contribution of the paper is threefold. First, it reframes sustainable healthcare AI as a knowledge-ecosystem problem rather than a narrow prediction problem. Second, it specifies governance requirements that make human-in-the-loop oversight meaningful rather than symbolic. Third, it offers a tractable blueprint for future institutional deployment in which governance precedes deployment, AI augments rather than displaces expert judgment, and sustainability is evaluated through interpretability, calibrated trust, workload, reproducibility, and organizational robustness as well as accuracy.

Keywords: Human-AI collaboration, Explainable clinical NLP, Healthcare governance, Socio-technical design, Knowledge management

Paper type: Academic Research Paper

1 Introduction

Narrative data occupy a structurally difficult but increasingly important position in healthcare organizations. Clinical notes, referral letters, discharge summaries, intake descriptions, and patient-written accounts contain knowledge that rarely fits neatly into structured fields. They capture sequencing, ambiguity, experiential nuance, and the language through which conditions are interpreted in practice. For this reason, they are often indispensable to professional judgment. At the same time, these same properties make narrative data hard to search, annotate, compare, and govern at scale.

This tension helps explain the appeal of AI and natural language processing in healthcare. If large volumes of free text can be transformed into usable signals, healthcare organizations may be better able to identify emerging risks, monitor trajectories, and support decision-making. Yet the dominant evaluative lens remains heavily accuracy-centered. Systems are commonly assessed in terms of benchmark performance, while the conditions that make them trustworthy in use remain under-specified. A model can be statistically impressive and still be opaque, context-bound, difficult to contest, or institutionally ungovernable (Topol, 2019; Obermeyer and Emanuel, 2016; Shortliffe and Sepulveda, 2018).

This paper addresses the following research question: what socio-technical and governance conditions are required for the sustainable development of human-in-the-loop AI systems based on health-related narrative data? It does so through PRAXIS, an early-stage doctoral research protocol designed from inception as human-in-the-loop. PRAXIS is presented here as an illustrative analytical case for deriving governance requirements, not

as a completed empirical study. Data collection is ongoing, no final results are reported, and the current phase does not constitute clinician-facing deployment.

The paper makes three contributions. First, it reframes sustainable healthcare AI as a knowledge-ecosystem problem rather than a narrow prediction problem. Second, it identifies design requirements that make human oversight meaningful rather than symbolic. Third, it introduces label fitting as a governance risk that arises when annotation schemes built on one population are transferred to others without explicit scope review. The argument proceeds by situating PRAXIS in the literature on healthcare AI, socio-technical systems, and knowledge management, then developing a governance-first framework organized around AI, human, and organizational knowledge-governance layers.

2 Background: From Accuracy to Sustainable Knowledge Ecosystems

2.1 AI in healthcare beyond accuracy

Early discussion of AI in healthcare often emphasized the possibility that machine-learning systems could match or exceed human performance on bounded tasks. That prospect remains important, but real-world implementation has exposed a persistent gap between benchmark results and practical reliability. Clinical impact depends not only on whether a system can classify, rank, or predict within a curated dataset, but also on whether its outputs are interpretable, whether the underlying data are fit for purpose, and whether the surrounding organization can absorb and govern the tool over time (Kelly et al., 2019; Cresswell et al., 2020).

This broader understanding has become central to trustworthy AI frameworks. Robustness, fairness, accountability, and interpretability matter because healthcare decisions are embedded in institutional settings where errors are costly and responsibilities cannot be externalized to the model (European Commission, 2019). The distinction between technical validity and practical usefulness is therefore crucial. A system may perform well under training conditions and still fail when users cannot understand what it is flagging, when the training population is too narrow, or when institutional routines do not support effective review (Obermeyer and Emanuel, 2016; Shortliffe and Sepulveda, 2018).

Natural language processing intensifies these concerns. When AI systems operate on narrative material rather than on structured variables, the link between outputs and source meaning is less transparent. Statistical regularities can easily be mistaken for validated clinical signals. In high-stakes contexts, this produces a recurring risk of overclaiming: the treatment of machine-generated associations as settled findings rather than as candidate signals requiring expert appraisal. Sustainable deployment therefore requires evaluative criteria that extend beyond accuracy to include interpretability, scope control, auditability, and equity across different populations.

2.2 Human-AI interaction as a socio-technical problem

Healthcare AI does not operate in a vacuum. It is embedded in socio-technical systems composed of professionals, routines, interfaces, institutional cultures, and accountability structures. Socio-technical research has long shown that technical outcomes depend on the arrangement of social practices around the tool and that social and technical elements are mutually constitutive (Carayon, 2006; Baxter and Sommerville, 2011). For narrative AI, this means that the usefulness of a model depends on who reviews outputs, what they are expected to do with them, and how disagreement or override is handled.

A well-documented danger in human-AI interaction is automation bias: the tendency to defer to automated outputs even when independent judgment would justify correction (Parasuraman and Manzey, 2010). In healthcare settings, this can erode the very quality of oversight that human-in-the-loop design is meant to protect. Equally problematic is the opposite condition, where system outputs are so poorly explained or so poorly integrated into work routines that they are ignored, distrusted, or treated as noise. In both cases, the challenge is organizational as much as technical.

For this reason, human-in-the-loop cannot be treated as a governance solution in itself. A human positioned at a decision point contributes meaningful oversight only when the work of interpretation has been explicitly designed. This requires clear review standards, documented reasons for accepting or rejecting suggestions, and protocols for handling uncertainty and disagreement (Amershi et al., 2019). In future clinician-facing applications, it also requires attention to workload, training, and role redesign so that AI augments rather than distorts professional judgment.

2.3 Knowledge governance, semantic entropy, and label fitting

A complementary perspective comes from knowledge management. From this viewpoint, corpora, annotation schemes, candidate labels, model outputs, and procedural documents are not merely technical artifacts. They are organizational knowledge assets whose value depends on the conditions under which they are created, validated, revised, and transferred (Davenport and Prusak, 1998; Nonaka and Takeuchi, 1995). If such assets are moved across settings without documented provenance or explicit scope control, the organization is not simply scaling knowledge. It is also scaling the hidden assumptions sedimented in those assets.

Recent work on the knowledge commoditization paradox sharpens this problem by arguing that the transformation of context-rich professional knowledge into AI-accessible representations can generate semantic entropy: a progressive loss of contextual information through computational abstraction. The issue is not that abstraction is always undesirable, but that it can strip away precisely the relational and situational cues that experts need in order to judge meaning appropriately. In this sense, AI-mediated representation is never a neutral translation process (Allega, Puzio and Rizzo, 2026b).

Companion work in other high-stakes domains suggests that trustworthy AI architectures should preserve provenance, explicit scope boundaries, and structured forms of contestability instead of collapsing all judgment into opaque outputs. Layered governance, auditable reasoning, and explicit management of disagreement have been proposed as ways to make hybrid systems more accountable over time (Allega, Puzio and Rizzo, 2026a; Allega, Puzio and Rizzo, 2026c; Allega, Puzio and Rizzo, 2026d). The present paper does not transfer legal AI content into healthcare. Rather, it adapts a governance logic from adjacent high-stakes domains to the analysis of health-related narratives.

Within this context, the paper introduces label fitting as a governance problem. Label fitting occurs when annotation schemes developed on a specific and narrow population encode interpretive assumptions that are locally reasonable but misleading when transferred elsewhere. It differs from technical distribution shift because the instability lies at the annotation and interpretation layer rather than only in feature distributions. Addressing label fitting therefore requires lifecycle governance, cross-population review, and explicit documentation of where an annotation scheme is valid and where it is not.

3 The PRAXIS Protocol as an Early-Stage Illustrative Case

PRAXIS is an early-stage doctoral research project aimed at developing governance-robust procedures for extracting structured signals from health-related narrative data produced in Italian. Its long-term horizon is translational: to support the systematic organizational use of narrative material in institutional and, potentially, clinical settings. In the present paper, however, PRAXIS functions as an early-stage case through which governance requirements can be made explicit before any strong claims about performance or clinical use are advanced.

The current phase involves a corpus of participant-generated Italian narratives collected from university students aged 19-24 through a standardized protocol. Participants produce free-text written responses designed to elicit personally relevant health-related accounts. This bounded setting is methodologically useful because it is ethically accessible and allows annotation procedures to be refined before extension to more sensitive populations. Data collection is ongoing and no final empirical results are available at this stage.

From inception, PRAXIS is human-in-the-loop. NLP and text-analysis components support signal extraction, linguistic structuring, and candidate annotation generation, but all interpretive judgments and final classification choices remain under expert human supervision. The annotation scheme is treated not as a finished technical product but as an evolving knowledge asset subject to revision, discrepancy review, and explicit documentation. The protocol is designed to extend to additional sub-populations, including clinical populations, in later phases. Clinician-facing deployment therefore remains a long-term translational objective rather than an empirical feature of the current phase.

4 A Governance-First Framework for Human-in-the-Loop AI

The framework proposed here organizes sustainable human-in-the-loop AI development around three interdependent layers: the AI layer, the human layer, and the organizational knowledge-governance layer. Each layer has a primary function, a characteristic set of risks, and a set of governance requirements. The key assumption is that sustainability cannot be achieved by strengthening only one layer. Risk in any layer propagates to the others, while governance at each layer presupposes adequate governance in the adjacent ones.

4.1 The AI layer

Within PRAXIS, the AI layer is responsible for signal extraction, linguistic structuring, and candidate annotation generation. It processes volumes of narrative material that would be analytically difficult to handle through unaided manual review alone and presents provisional outputs that expert reviewers can evaluate, modify, or reject. The value of this layer lies in amplification and ordering, not in autonomous decision-making. In the current phase, its role is therefore explicitly supportive and bounded.

The principal governance risks at this layer are overclaiming, opacity, overfitting, and boundary violation. Overclaiming occurs when machine-generated patterns are treated as established findings rather than as candidate signals. Opacity makes it difficult for reviewers to understand why a suggestion has been produced. Overfitting arises when models or procedures reflect a narrow construction context too closely. Boundary violation occurs when outputs are applied outside the population, domain, or task conditions for which they were developed. These are not marginal defects; they determine whether later human review can be meaningful at all.

Effective governance of the AI layer therefore requires interpretability standards, error-analysis routines, versioned change control, and explicit scope documentation. In more mature institutional settings, such standards may be operationalized through evidence spans, uncertainty cues, model cards, or privacy-preserving data architectures. Yet the governance point is prior to any specific method: AI outputs must remain inspectable, documented, and contestable, and any change to model components or annotation suggestions must be traceable over time (Mitchell et al., 2019; European Commission, 2019).

4.2 The human layer

In the current PRAXIS phase, the human layer is occupied by expert reviewers rather than clinicians. In later translational phases, the same governance logic would extend to clinician-facing use. The substantive point is that human involvement is not a symbolic checkpoint. It is a knowledge practice through which candidate signals are interpreted, contextual meaning is restored, ambiguous cases are adjudicated, and the limits of the category system are made visible. Final responsibility therefore remains human even when AI contributes to the organization of the evidence.

This layer is vulnerable to automation bias, interpretive inconsistency, reasoning-traceability loss, and excessive reliance on already stabilized categories. Reviewers may defer too quickly to machine suggestions, disagree without documented rationale, or overlook cases that do not fit the current annotation vocabulary. In a healthcare context, these failures are serious because they affect not only immediate classification quality but also the long-term calibration of the system and the organization's confidence in it (Parasuraman and Manzey, 2010; Amershi et al., 2019).

Governance of the human layer requires explicit annotation guidelines, co-annotation and discrepancy-review practices, protocols for calibrated reliance, and documentation of interpretive rationale. Disagreement is especially important. Rather than being treated only as noise to be minimized, it should also be treated as information about the adequacy of the scheme and the uncertainty of the underlying case. This makes the human layer both a review mechanism and a learning mechanism for the broader knowledge ecosystem (Artstein and Poesio, 2008; Allega, Puzio and Rizzo, 2026d).

4.3 The organizational knowledge-governance layer

The organizational layer treats the corpus, annotation schemes, output histories, and procedural documents as maintained knowledge assets rather than as disposable project by-products. It is here that provenance, access, revision, and transferability are governed. This layer matters because even well-designed AI tools and well-trained reviewers can produce fragile systems if the surrounding organization does not record how assets were built, under what conditions they can be reused, and how they should be revised over time (Davenport and Prusak, 1998; Nonaka and Takeuchi, 1995).

The central risk foregrounded in this paper - label fitting - sits at this layer. When labels built on one population are reused in another without structured review, the organization can mistake local adequacy for general validity. The result is not simply lower performance. It is a loss of semantic integrity. In effect, the organization begins to treat historically contingent annotation choices as if they were stable knowledge. This risk is amplified when documentation is fragmented, when provenance is weak, or when asset reuse is driven by convenience rather than governed scope control (Allega, Puzio and Rizzo, 2026b).

Governance at the organizational layer therefore requires four minimum capabilities: provenance documentation for corpus elements and protocol versions; access governance defining who can use, modify, or

redistribute assets; lifecycle governance specifying how assets are reviewed, revised, and retired; and cross-population review procedures that test transferability before extension to new groups. In later institutional settings, these routines can also support organizational learning loops in which frontline review feeds back into training, threshold setting, and documentation updates. Table 1 summarizes the framework.

Table 1: Governance-first framework for sustainable human-AI narrative analysis

Layer	Key risks	Governance requirements
AI	Overclaiming Opacity Overfitting Boundary violation	Interpretability standards Error analysis Versioning Scope documentation
Human	Automation bias Interpretive inconsistency Reasoning-traceability loss Category lock-in	Annotation guidelines Co-annotation and discrepancy review Calibrated reliance protocols Documentation of interpretive rationale
Organizational knowledge governance	Fragmentation Weak reproducibility Label fitting Misapplication	Provenance documentation Access governance Lifecycle governance Cross-population review protocols

Taken together, the three layers show why sustainable healthcare AI cannot be reduced to a question of model quality. The AI layer must be interpretable and scoped; the human layer must be structured and traceable; and the organizational layer must preserve provenance and manage transferability. Sustainability is achieved only when these layers are aligned and jointly maintained.

5 Discussion

5.1 Sustainability as a property of the whole configuration

The main implication of the framework is that sustainable AI in healthcare is a property of the socio-technical configuration as a whole, not of the AI component in isolation. This has two consequences. First, evaluation must move beyond predictive metrics toward a broader assessment of interpretability, calibrated trust, workload, reproducibility, workflow compatibility, governance compliance, and equity. Second, the relevant organizational question is not simply whether a model is accurate, but whether the organization has built the human and governance capacities required to use its outputs responsibly (Baxter and Sommerville, 2011; European Commission, 2019).

This shift also reframes human capability. In a governance-first ecosystem, valuable human capabilities include not only domain expertise but also the ability to contest outputs, document rationale, calibrate reliance, manage disagreement, and steward evolving knowledge assets. These are organizational capabilities as much as individual ones. They require training, documented routines, and role clarity. In future clinician-facing settings, the clinician is not displaced by AI; rather, the clinician becomes more explicitly an accountable interpreter of machine-assisted signals embedded in governed workflows.

The framework likewise suggests that the most useful contribution of AI to healthcare organizations may be organizational rather than purely predictive. By making narrative material more tractable for review, AI can help convert diffuse textual material into maintained knowledge resources. But that conversion is sustainable only if semantic indicators remain linked to provenance, uncertainty, and revision histories. Otherwise, organizations risk institutionalizing decontextualized proxies that are easier to process but less meaningful in practice.

5.2 Practical implications and limitations

Practically, the framework encourages organizations to invest early in governance before attempting broad deployment. In the PRAXIS case, this means treating annotation design, discrepancy logging, and scope documentation as core development tasks rather than as administrative add-ons. The same logic can guide future institutional work: governance should begin at the moment data are collected and categories are created, not after a model has already been optimized. In this sense, the framework offers a blueprint for organizations that want AI to augment professional knowledge without eroding accountability.

The paper also clarifies why an early-stage non-clinical protocol can be analytically valuable. Working first with an ethically accessible student corpus does not solve the challenges of clinical deployment, but it does make

visible the conditions under which interpretive schemes are built, revised, and stabilized. This is precisely where hidden assumptions can become sedimented and later misapplied. The current phase therefore provides a controlled environment for identifying governance requirements before institutional use expands.

At the same time, the paper has clear limitations. The framework is conceptual and analytically derived rather than empirically validated. It has not yet been tested against completed performance results, multi-site deployment, or practitioner workflow studies. The student corpus is useful for protocol development but cannot substitute for work with clinical populations. Future research should therefore complete the current data-collection phase, examine how label fitting appears across additional sub-populations, and test the framework with institutional actors, clinician users, and longitudinal governance measures

6 Conclusions

This paper has argued that sustainable AI development in healthcare begins before deployment, in the governed construction of data, labels, review practices, and organizational conditions. Using PRAXIS as an early-stage illustrative case, it has proposed a governance-first framework organized around AI, human, and organizational knowledge-governance layers. The framework treats narrative AI not as a stand-alone predictive technology but as part of a broader human-AI knowledge ecosystem whose sustainability depends on provenance, interpretability, calibrated reliance, and controlled transfer across contexts.

The central governance problem emphasized here is label fitting: the embedding of population-specific interpretive assumptions in annotation schemes that are later transferred without adequate scope review. Framed this way, sustainable healthcare AI is not achieved by accuracy alone. It requires institutions that can preserve semantic integrity, document uncertainty, and keep human accountability active as systems evolve. PRAXIS remains early-stage, and the framework now awaits empirical testing. Its contribution is to specify, in operational terms, what governance-first AI development requires and why that requirement is foundational rather than optional.

7 Ethics Declaration

This paper discusses an ongoing research protocol (PRAXIS) involving participant-generated written narratives from consenting adults aged 19-24. The manuscript does not report completed clinical research results and does not involve personal health data from clinical populations in the current phase. The protocol is conducted in accordance with applicable ethical principles for research involving human participants, including informed consent, data anonymization, confidentiality, and relevant data-protection requirements. Any future extension to clinical populations or institutional deployment will require review and authorization under the relevant procedures before implementation.

8 AI Declaration

AI-assisted writing tools were used for outlining, language refinement, and editorial revision during the preparation of this manuscript. They were not used to generate empirical data, statistical findings, or unverified references. All arguments, factual claims, citationsand bibliographic details were reviewed and approved by the authors, who retain full responsibility for the content of the paper.

References

Allega, D., Puzio, G. and Rizzo, G. (2026a) 'The Artificial Legal Intelligence Society (ALIS) as an open, multi-sided platform for law-as-computation', in Panait, M., Radulescu, I.G., Tudorica, B., Popescu, C., Voica, M.C. and Nica, M. (eds.) Proceedings of the 4th International Conference Creativity and Innovation in Digital Economy. Ploiesti: Petroleum-Gas University of Ploiesti Publishing House, pp. 66-75.

Allega, D., Puzio, G. and Rizzo, G. (2026b) 'The knowledge commoditization paradox: Theoretical and practical challenges of AI-driven value extraction in information-intensive organizations', in Panait, M., Radulescu, I.G., Tudorica, B., Popescu, C., Voica, M.C. and Nica, M. (eds.) Proceedings of the 4th International Conference Creativity and Innovation in Digital Economy. Ploiesti: Petroleum-Gas University of Ploiesti Publishing House, pp. 95-107.

Allega, D., Puzio, G. and Rizzo, G. (2026c) 'MERL-T: A multi-expert architecture for trustworthy artificial legal intelligence', in Panait, M., Radulescu, I.G., Tudorica, B., Popescu, C., Voica, M.C. and Nica, M. (eds.) Proceedings of the 4th International Conference Creativity and Innovation in Digital Economy. Ploiesti: Petroleum-Gas University of Ploiesti Publishing House, pp. 76-83.

Allega, D., Puzio, G. and Rizzo, G. (2026d) 'Reinforcement learning from community feedback: A novel framework for artificial intelligence in social science domains', in Panait, M., Radulescu, I.G., Tudorica, B., Popescu, C., Voica, M.C. and Nica, M.

(eds.) Proceedings of the 4th International Conference Creativity and Innovation in Digital Economy. Ploiesti: Petroleum-Gas University of Ploiesti Publishing House, pp. 84-93.

Amershi, S., Weld, D., Vorvoreanu, M. et al. (2019) 'Guidelines for human-AI interaction', Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, pp. 1-13.

Artstein, R. and Poesio, M. (2008) 'Inter-coder agreement for computational linguistics', Computational Linguistics, 34(4), pp. 555-596.

Baxter, G. and Sommerville, I. (2011) 'Socio-technical systems: From design methods to systems engineering', Interacting with Computers, 23(1), pp. 4-17.

Carayon, P. (2006) 'Human factors of complex sociotechnical systems', Applied Ergonomics, 37(4), pp. 525-535.

Cresswell, K., Callaghan, M., Khan, S. et al. (2020) 'Investigating the use of data-driven artificial intelligence in computerised decision support systems for health and social care: A systematic review', Health Informatics Journal, 26(3), pp. 2138-2147.

Davenport, T.H. and Prusak, L. (1998) Working Knowledge: How Organizations Manage What They Know. Boston: Harvard Business School Press.

European Commission (2019) Ethics Guidelines for Trustworthy AI. Brussels: High-Level Expert Group on Artificial Intelligence.

Kelly, C.J., Karthikesalingam, A., Suleyman, M. et al. (2019) 'Key challenges for delivering clinical impact with artificial intelligence', BMC Medicine, 17, 195.

Mitchell, M., Wu, S., Zaldivar, A. et al. (2019) 'Model cards for model reporting', Proceedings of the Conference on Fairness, Accountability, and Transparency, pp. 220-229.

Nonaka, I. and Takeuchi, H. (1995) The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation. New York: Oxford University Press.

Obermeyer, Z. and Emanuel, E.J. (2016) 'Predicting the future - Big data, machine learning, and clinical medicine', New England Journal of Medicine, 375(13), pp. 1216-1219.

Parasuraman, R. and Manzey, D.H. (2010) 'Complacency and bias in human use of automation: An attentional integration', Human Factors, 52(3), pp. 381-410.

Shortliffe, E.H. and Sepulveda, M.J. (2018) 'Clinical decision support in the era of artificial intelligence', JAMA, 320(21), pp. 2199-2200.

Topol, E.J. (2019) 'High-performance medicine: The convergence of human and artificial intelligence', Nature Medicine, 25(1), pp. 44-56.

From Paradoxes and Entropies to Sustainable Knowledge Ecosystems: Architectural and Governance Principles

Daniele Allega[1], Gabriele Rizzo[2]
[1] Universitas Mercatorum, Rome, Italy
[2] LUISS Guido Carli University, Rome, Italy

Abstract

Artificial Intelligence is often presented as a shortcut to organizational sustainability, yet many AI initiatives underperform because firms attempt to automate before they have structured, governed, and shared their knowledge assets. This paper addresses that gap by asking how organizations can move from fragmented, tool-centred AI adoption to sustainable knowledge ecosystems. We build a design-oriented conceptual framework through analytical generalization from four design artifacts previously developed in the legal-AI domain: the knowledge commoditization paradox (Allega, Puzio and Rizzo, 2026b), the MERL-T multi-expert architecture (Allega, Puzio and Rizzo, 2026c), the RLCF governance model (Allega, Puzio and Rizzo, 2026d), and the ALIS platform framework (Allega, Puzio and Rizzo, 2026a). The legal domain is treated as a revelatory setting because it concentrates stringent requirements for traceability, pluralism, accountability, and knowledge integrity that are increasingly relevant across knowledge-intensive organizations. The paper makes two main contributions. First, it extends the earlier notion of semantic entropy by identifying three entropic pressures that undermine the Knowledge Journey: coordination entropy, generated by fragmented repositories and inconsistent taxonomies; semantic entropy, generated by the decontextualization of knowledge during computational representation; and governance entropy, generated by opaque AI-mediated workflows that dilute authority and responsibility. Second, it translates these pressures into five design principles for sustainable knowledge ecosystems: commons-first knowledge architecture, provenance-preserving services, epistemically differentiated human-AI collaboration, community-anchored governance, and sustainability-by-measurement. These principles are then mapped onto the three phases of Planning, Acquisition, and Impact to show how organizations can reduce hidden costs such as search time, duplication of effort, hallucinated outputs, accountability gaps, and knowledge lock-in. Rather than treating AI as a stand-alone capability, the paper argues that sustainable value emerges when architectural choices and governance mechanisms are designed together. The result is a practical and theoretically grounded model for organizations seeking to transform knowledge management from a hidden cost centre into a resilient, auditable, and sustainability-oriented ecosystem.

Keywords: Knowledge management; Knowledge ecosystems; AI governance; Human-AI collaboration; Corporate sustainability.

Paper type: Academic Research Paper

1 Introduction

Artificial Intelligence has become a strategic horizon for organizations seeking resilience, efficiency, and sustainability. Yet, in practice, many AI programs begin at the wrong point. They start from models, copilots, or chat interfaces, while the knowledge infrastructure on which those tools depend remains fragmented, weakly governed, and poorly documented. As a result, organizations often accelerate access to inconsistent content rather than improving the quality of organizational knowing. The hidden costs are familiar: time lost searching for information, duplicated work, obsolete versions of documents, repeated reinvention of solutions, and growing dependence on opaque external platforms.

This mismatch reveals a broader conceptual problem. In many organizational debates, technological capability is treated as if it were equivalent to strategic readiness. However, knowledge-intensive work depends not only on access to information, but also on the ability to preserve context, identify provenance, recognize expertise, and govern the interaction between human judgement and machine outputs. Classical knowledge management already emphasized that valuable knowledge is not exhausted by codified information (Polanyi, 1966; Nonaka and Takeuchi, 1995). More recent research shows that AI can also reconfigure professional practice, redistribute authority, and create new forms of opacity and dependence (Pasquale, 2015; Pachidi et al., 2021; Susskind and Susskind, 2015).

The present paper addresses this gap through a design-oriented perspective. We draw on a research program originally developed in the legal-AI domain, where requirements for traceability, accountability, and interpretive plurality are particularly demanding. That work produced four complementary artifacts: a theoretical account of the knowledge commoditization paradox (Allega, Puzio and Rizzo, 2026b), a multi-expert architecture for trustworthy reasoning (Allega, Puzio and Rizzo, 2026c), a governance model for community-based validation

(Allega, Puzio and Rizzo, 2026d), and a platform framework that combines commons and differentiated services (Allega, Puzio and Rizzo, 2026a). Although domain-specific in origin, these artifacts expose more general organizational problems that are increasingly visible across industries.

The paper therefore asks two questions. First, how can the paradoxes created by AI-mediated knowledge management be reorganized into actionable design problems? Second, which architectural and governance principles can guide the Knowledge Journey toward sustainable value creation? To answer them, we extend the earlier notion of semantic entropy into a broader framework of organizational entropies and derive five principles for sustainable knowledge ecosystems.

The contribution is threefold. Theoretically, the paper reframes AI-enabled knowledge management as an ecosystem design problem rather than a tool-adoption problem. Conceptually, it introduces coordination entropy and governance entropy alongside semantic entropy. Practically, it offers an operational model aligned with the three phases of Planning, Acquisition, and Impact showing how firms can turn hidden costs into sustainable competitive advantage.

2 Theoretical Background

2.1 The knowledge commoditization paradox

The starting point of our argument is the knowledge commoditization paradox: the simultaneous necessity and impossibility of reducing complex organizational knowledge to computationally accessible and tradable representations (Allega, Puzio and Rizzo, 2026b). AI systems require knowledge to be structured, indexed, embedded, linked, or otherwise made processable. Yet the very operations that increase computational accessibility tend to strip away the contextual richness that makes knowledge valuable in organizational settings. This tension is visible in the movement from documents and practices to vectors, metadata, and machine-readable schemas.

This problem has both economic and epistemic roots. Arrow (1962) showed that information markets are marked by deep asymmetries: value is difficult to assess before use, yet once revealed it becomes difficult to appropriate. In organizational settings, the challenge is amplified because knowledge is not a neutral object. It is embedded in routines, interpretive communities, hierarchies of authority, and situated practices. When organizations try to externalize that knowledge for AI systems, they face a difficult choice between oversimplified codification and increasingly expensive representational complexity.

Earlier work in our research stream described this process through the concept of semantic entropy: the progressive loss of contextual information through computational abstraction (Allega, Puzio and Rizzo, 2026b). This diagnosis is consistent with broader critiques of statistical language systems that capture form and correlation more easily than grounded meaning and pragmatic context (Bender and Koller, 2020). It is also reinforced by recent research on retrieval architectures. Weller et al. (2025) show that single-vector embedding systems face structural limits in representing all possible relevance combinations, suggesting that some knowledge access failures are not merely training-data problems but architectural bottlenecks.

The implication is significant for knowledge management. If computational representations are intrinsically selective, then organizations cannot assume that more AI automatically means better knowledge. They must instead decide what kind of knowledge losses are acceptable, what forms of human intervention remain necessary, and what architectural choices can reduce avoidable distortions.

2.2 From semantic entropy to coordination and governance entropy

We extend the previous concept of semantic entropy by proposing two additional entropic pressures that affect the organizational Knowledge Journey. We use entropy analogically but operationally, following the intuition that systems lose coherence when structure, signal, and accountability dissipate across transformations (Shannon, 1948).

Coordination entropy refers to the dissipation generated by fragmented repositories, inconsistent taxonomies, poorly maintained metadata, and misaligned workflows. Here the problem is not only that knowledge is reduced during computational processing, but that organizations cannot even coordinate what counts as relevant knowledge, where it resides, who owns it, or how it should circulate. Coordination entropy manifests in duplicated work, search costs, onboarding delays, and recurrent reconstruction of prior knowledge.

Governance entropy refers to the dilution of authority, responsibility, and legitimacy across opaque AI-mediated processes. As knowledge moves through retrieval tools, automated classifications, external APIs, and generative interfaces, the chain of accountability becomes harder to reconstruct. Who validated the source?

Who designed the ranking logic? Who is responsible for errors, omissions, bias, or overconfident outputs? In this sense, governance entropy is not merely a legal or ethical issue; it is an organizational knowledge problem because it affects whether outputs can be trusted, contested, and improved.

These three entropies are recursively related. Poor coordination amplifies semantic loss because inconsistent architectures increase the risk that models retrieve the wrong representations. Semantic loss amplifies governance problems because opaque outputs are harder to audit when context has already been stripped away. Governance failures, in turn, worsen coordination because actors lose incentives to share, curate, and maintain common knowledge resources. The Knowledge Journey is therefore destabilized not by a single defect, but by an entropic cascade.

2.3 Knowledge ecosystems as hybrid and governed infrastructures

If AI-mediated knowledge management is shaped by entropic pressures, then sustainable response requires more than better interfaces or larger models. It requires organizations to think in terms of knowledge ecosystems: layered socio-technical arrangements in which human expertise, codified resources, computational services, and governance mechanisms are designed together.

This ecosystem view shifts the focus from isolated tools to infrastructural choices. Platform research has long shown that control over standards, access points, and boundary resources structures who can participate in value creation and on what terms (Parker, Van Alstyne and Choudary, 2016). In knowledge-intensive settings, the same logic applies to ontologies, APIs, retrieval services, and validation rules. Without explicit governance, firms risk reproducing new asymmetries of dependency and opacity (Pasquale, 2015; Varoufakis, 2023).

Our prior work responds to this challenge by combining a commons layer, provenance-aware services, community-based governance, and differentiated applications (Allega, Puzio and Rizzo, 2026a). The broader organizational lesson is that sustainable knowledge ecosystems are hybrid by design. They do not eliminate human judgement, nor do they reject automation. Rather, they distribute roles between humans and machines in ways that preserve knowledge integrity while enabling scale, reuse, and continuous improvement.

3 Research Design

This paper is an academic conceptual study informed by design-science research (Hevner et al., 2004). Its purpose is not to test a single causal hypothesis, but to derive prescriptive design knowledge from a coherent set of previously developed artifacts. The analytical strategy is one of conceptual synthesis and analytical generalization from a revelatory domain: legal AI.

The legal domain is particularly suitable for this role because it concentrates requirements that are increasingly relevant across knowledge-intensive organizations: traceability of sources, explicit reasoning paths, interpretive plurality, temporal versioning, and high accountability for error. In other words, it exposes at high intensity the same tensions that other organizations face when deploying AI over strategic knowledge assets.

The empirical and conceptual inputs used for the synthesis are four artifacts reported in the CIDE 2025 proceedings. The first identified the knowledge commoditization paradox and formalized semantic entropy as a degradation function (Allega, Puzio and Rizzo, 2026b). The second developed MERL-T, a multi-expert architecture that combines explicit knowledge representation, differentiated reasoning, and provenance-preserving retrieval (Allega, Puzio and Rizzo, 2026c). The third introduced RLCF, a governance framework based on dynamic authority, uncertainty-preserving aggregation, and bias-aware community validation (Allega, Puzio and Rizzo, 2026d). The fourth integrated these ideas into a platform architecture structured around commons, services, governance, and applications (Allega, Puzio and Rizzo, 2026a).

The synthesis proceeded in four steps. First, we identified recurring problems across the four artifacts. Second, we clustered these problems into three higher-order entropic pressures: coordination, semantic, and governance entropy. Third, we mapped the design responses developed in the legal domain to those pressures. Fourth, we generalized them into organizational principles that can guide the Knowledge Journey across the phases of Planning, Acquisition, and Impact.

The resulting framework should therefore be read as design knowledge with both theoretical and managerial intent: it does not claim universal optimality, but it specifies a coherent set of architectural and governance principles that organizations can adapt to their own sectoral and strategic contexts.

Table 1: Entropic pressures, organizational manifestations, and design responses

Entropic pressure	Typical hidden cost	Primary design response	Indicative metrics
Coordination entropy	Search time, duplication, obsolete versions, siloed know-how	Commons-first knowledge architecture	Search time, reuse rate, onboarding time
Semantic entropy	Context loss, brittle retrieval, hallucinated outputs, weak explainability	Provenance-preserving services and differentiated human-AI collaboration	Provenance coverage, expert override rate, retrieval quality
Governance entropy	Accountability gaps, bias, opaque decisions, platform lock-in	Community-anchored governance and sustainability-by-measurement	Auditability, participation diversity, incident rate

4 Architectural and Governance Principles for Sustainable Knowledge Ecosystems

4.1 Planning phase: commons-first knowledge architecture

The first principle is that organizations should design a knowledge commons before they scale AI applications. In practice, this means investing in explicit taxonomies, metadata conventions, document ownership rules, versioning logic, and shared schemas for critical knowledge domains. The purpose is not bureaucratic neatness. It is to reduce coordination entropy by creating a stable layer of organizational meaning on top of which computational services can operate.

Our prior legal-domain implementations illustrate why this matters. Knowledge graphs, argument structures, and typed entities were not introduced as academic embellishments, but as mechanisms for making relations explicit, queryable, and governable (Allega, Puzio and Rizzo, 2026a; 2026c). The same logic generalizes to other sectors. A company cannot expect an intelligent intranet, an AI copilot, or a semantic search system to produce robust results if basic categories, ownership responsibilities, and update rules remain unstable.

A commons-first architecture also has sustainability implications. It reduces redundant local solutions, limits dependence on proprietary vendor logic, and creates reusable assets that can support multiple applications over time. From a knowledge-based view of the firm, such shared and governable knowledge resources are not peripheral infrastructure but a basis for sustained advantage because they improve the integration and application of specialized knowledge (Grant, 1996). This is particularly important in organizations where knowledge turnover, mergers, and distributed work routinely degrade continuity. Leadership therefore matters not only as sponsorship, but as stewardship: someone must be responsible for maintaining the architecture of organizational knowledge, not merely for purchasing the next AI tool.

4.2 Acquisition phase: provenance-preserving human-AI integration

The second and third principles address the Acquisition phase. First, knowledge services should preserve provenance. Retrieval, summarization, recommendation, and reasoning systems should expose their sources, timestamps, confidence signals, and evidentiary paths rather than producing answer-only outputs. Provenance is the organizational antidote to semantic entropy because it reconnects machine-readable representations to the contexts from which they were derived.

Second, organizations should design human-AI collaboration in an epistemically differentiated way. Tacit knowledge cannot be captured by indiscriminate ingestion alone. It emerges through interviews, process documentation, exception handling, narratives of practice, and expert interpretation. For this reason, AI should not be configured as a monolithic substitute for expertise. It should be embedded in workflows that distinguish between types of knowledge and types of reasoning.

The MERL-T artifact is instructive here. Its contribution was not simply better performance, but the separation of interpretive roles into differentiated expert modules combined through explicit synthesis logic (Allega, Puzio and Rizzo, 2026c). The broader organizational lesson is that sustainable knowledge acquisition requires role clarity. Rule-based content, conceptual definitions, experiential judgement, and precedent-like cases often demand different treatment. When everything is collapsed into a single black-box interaction, organizations gain superficial fluency but lose the ability to understand where knowledge came from, why it was selected, and when a human expert should intervene.

This point also resonates with the external literature on retrieval limits. If certain failures of knowledge access are architectural rather than purely statistical, then organizations should resist the temptation to treat larger

embeddings or more data as universal remedies (Weller et al., 2025). Hybrid approaches - combining structured repositories, explicit links, retrieval services, and human validation - are better suited to preserving the contextual integrity of strategic knowledge.

In practical terms, the Acquisition phase should therefore include structured elicitation routines, provenance-aware APIs, and expert review checkpoints. AI can accelerate extraction and integration, but it should do so in ways that preserve human experience and intuition as first-class organizational resources rather than residual corrections.

4.3 Impact phase: contestable governance and sustainability-by-measurement

The fourth principle is that sustainable knowledge ecosystems require contestable governance. Once knowledge is collected and distributed through AI-mediated services, organizations need mechanisms that specify who can contribute, who can validate, how disagreement is handled, and how risk is escalated. Governance cannot be postponed to compliance departments after deployment; it must be embedded into the ecosystem itself.

The RLCF and ALIS artifacts point to four useful mechanisms in this regard: dynamic authority, uncertainty-preserving aggregation, explicit transparency rules, and structured challenge functions such as Devil's Advocate roles (Allega, Puzio and Rizzo, 2026a; 2026d). These mechanisms matter because not all knowledge domains converge on a single correct answer, and organizations often damage learning when they force premature consensus. In high-stakes environments, responsible AI should sometimes present bounded disagreement rather than spurious certainty. This is consistent with recent alignment research arguing that pluralistic AI should preserve heterogeneous perspectives rather than collapse them into a single dominant reward signal (Sorensen et al., 2024; Chakraborty et al., 2024).

The fifth principle follows directly: sustainability must be measured through feedback loops, not assumed as a by-product of adoption. Many organizations evaluate AI systems through local efficiency metrics only, such as faster response times or reduced service tickets. Those indicators matter, but they do not reveal whether the knowledge ecosystem is becoming more robust. A sustainability-oriented measurement system should therefore include both operational and epistemic indicators: search time, reuse rate, provenance coverage, expert override rate, auditability of outputs, participation diversity in validation, decision latency for ESG-related issues, and the persistence of organizational knowledge despite turnover.

This broader view of impact also aligns with emerging expectations around trustworthy and accountable AI, including requirements for transparency, human oversight, and traceability in high-risk contexts (European Union, 2024). More importantly, it recognizes that organizational sustainability is not only environmental. It is also cognitive and institutional. A sustainable knowledge ecosystem is one that preserves continuity, supports informed decisions, and can be corrected when its classifications, recommendations, or summaries are found wanting.

5 Discussion

The framework developed in this paper contributes to knowledge management research by shifting the unit of analysis from isolated AI tools to sustainable knowledge ecosystems. Much current debate still assumes that the core managerial challenge is adoption: whether firms can introduce copilots, retrieval tools, or generative systems into existing processes. Our findings suggest that this view is too narrow. The decisive issue is architectural and governance readiness. Organizations do not fail only because models are inaccurate; they fail because the ecosystems into which those models are inserted are entropically fragile.

A first theoretical contribution is therefore the extension of semantic entropy into a wider three-part model. Semantic entropy captures loss of context during computational representation, but organizations also suffer from coordination entropy and governance entropy. This broader framework helps explain why AI programs can disappoint even when the underlying model appears technically strong. In this sense, the problem is not simply one of information processing, but of preserving and organizing multiple forms of knowledge under conditions of pluralism, which recent alignment research has likewise identified as a central challenge for AI systems intended to serve heterogeneous communities (Sorensen et al., 2024; Chakraborty et al., 2024). Poor taxonomy design, missing ownership, fragmented repositories, and opaque decision chains can neutralize or distort otherwise useful systems.

A second contribution is methodological and design-oriented. By analytically generalizing from a high-accountability domain, the paper shows how artifact-based research can inform broader KM theory. The legal field is not identical to manufacturing, healthcare, consulting, or the public sector, but it makes visible a set of

design demands - provenance, contestability, explicit reasoning, pluralism - that are becoming increasingly relevant across knowledge-intensive organizations.

A third contribution concerns sustainability. In many organizational discussions, sustainability enters the AI agenda late, often as a reporting or compliance category. The present framework suggests that sustainability begins much earlier, in how knowledge is structured, governed, and renewed. Reducing search waste, limiting duplication, preserving institutional memory, avoiding hallucinated outputs, and making decisions auditable are not secondary efficiency gains. They are part of building organizations that can endure change without losing the quality of their knowing.

For managers, the paper implies a sequence of intervention. First, map where coordination entropy is highest: fragmented repositories, inconsistent tags, undocumented expert routines. Second, build a commons layer with explicit ownership and versioning rules. Third, ensure that AI services expose provenance and that human expertise is segmented rather than flattened. Fourth, embed governance mechanisms that preserve contestability and learning. Fifth, measure the ecosystem through both efficiency and epistemic integrity indicators.

The framework also has clear limits. It is conceptual and design-oriented, not a large-sample empirical validation. The proposed principles emerged from one revelatory domain and should now be tested across sectors such as manufacturing, healthcare, and public administration. In addition, organizations differ in scale and maturity; not all need full knowledge graphs or elaborate governance boards. The relevant claim is therefore not that one architecture fits all, but that sustainable AI requires some explicit combination of commons, provenance, differentiated collaboration, governance, and feedback.

6 Conclusion

This paper argued that the path from AI adoption to organizational sustainability passes through the deliberate design of knowledge ecosystems. Starting from the knowledge commoditization paradox, we identified an entropic cascade - coordination, semantic, and governance entropy - that helps explain why many AI initiatives struggle to create durable value. We then derived five principles to counter those pressures: commons-first knowledge architecture, provenance-preserving services, epistemically differentiated human-AI collaboration, community-anchored governance, and sustainability-by-measurement.

The central implication is straightforward. Organizations should stop treating AI as an isolated capability and start treating it as a knowledge-infrastructure challenge. Sustainable value does not emerge when a model is merely added to existing processes. It emerges when knowledge assets, human expertise, and governance mechanisms are intentionally configured as a resilient ecosystem. Under these conditions, the hidden costs of fragmented knowledge can be transformed into a durable source of organizational learning, accountability, and sustainable advantage.

7 Ethics Declaration

This paper is conceptual and relies on published sources and previously reported design artifacts. It does not involve human participants, personal data collection, or experimental intervention. Consequently, formal ethical clearance was not required.

8 AI Declaration

An AI language tool was used to support drafting, language refinement, and formatting under full human supervision. The tool was not used to fabricate evidence or citations. All arguments, source selection, and final editorial decisions were reviewed and validated by the authors.

References

Allega, D., Puzio, G. and Rizzo, G. (2026a) 'The Artificial Legal Intelligence Society (ALIS) as an Open, Multi-Sided Platform for Law-as-Computation', in Panait, M., Rădulescu, I.G., Tudorică, B., Popescu, C., Voica, M.C. and Nica, M. (eds.) Proceedings of the 4th International Conference Creativity and Innovation in Digital Economy. Ploiești: Petroleum-Gas University of Ploiești Publishing House, pp. 66–75.

Allega, D., Puzio, G. and Rizzo, G. (2026b) 'The Knowledge Commoditization Paradox: Theoretical and Practical Challenges of AI-Driven Value Extraction in Information-Intensive Organizations', in Panait, M., Rădulescu, I.G., Tudorică, B., Popescu, C., Voica, M.C. and Nica, M. (eds.) Proceedings of the 4th International Conference Creativity and Innovation in Digital Economy. Ploiești: Petroleum-Gas University of Ploiești Publishing House, pp. 95–107.

Allega, D., Puzio, G. and Rizzo, G. (2026c) 'MERL-T: A Multi-Expert Architecture for Trustworthy Artificial Legal Intelligence', in Panait, M., Rădulescu, I.G., Tudorică, B., Popescu, C., Voica, M.C. and Nica, M. (eds.) Proceedings of the 4th International Conference Creativity and Innovation in Digital Economy. Ploiești: Petroleum-Gas University of Ploiești Publishing House, pp. 76–83.

Allega, D., Puzio, G. and Rizzo, G. (2026d) 'Reinforcement Learning from Community Feedback: A Novel Framework for Artificial Intelligence in Social Science Domains', in Panait, M., Rădulescu, I.G., Tudorică, B., Popescu, C., Voica, M.C. and Nica, M. (eds.) Proceedings of the 4th International Conference Creativity and Innovation in Digital Economy. Ploiești: Petroleum-Gas University of Ploiești Publishing House, pp. 84–93.

Arrow, K.J. (1962) 'Economic Welfare and the Allocation of Resources for Invention', in The Rate and Direction of Inventive Activity: Economic and Social Factors. Princeton, NJ: Princeton University Press, pp. 609–626.

Bender, E.M. and Koller, A. (2020) 'Climbing towards NLU: On Meaning, Form, and Understanding in the Age of Data', in Jurafsky, D., Chai, J., Schluter, N. and Tetreault, J. (eds.) Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. Online: Association for Computational Linguistics, pp. 5185–5198.

Chakraborty, S., Qiu, J., Yuan, H., Koppel, A., Manocha, D., Huang, F., Bedi, A. and Wang, M. (2024) 'MaxMin-RLHF: Alignment with Diverse Human Preferences', in Proceedings of the 41st International Conference on Machine Learning. Proceedings of Machine Learning Research, 235, pp. 6116–6135.

European Union (2024) Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Official Journal of the European Union, OJ L, 2024/1689, 12 July.

Grant, R.M. (1996) 'Toward a Knowledge-Based Theory of the Firm', Strategic Management Journal, 17(S2), pp. 109–122.

Hevner, A.R., March, S.T., Park, J. and Ram, S. (2004) 'Design Science in Information Systems Research', MIS Quarterly, 28(1), pp. 75–105.

Nonaka, I. and Takeuchi, H. (1995) The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation. Oxford: Oxford University Press.

Pachidi, S., Berends, H., Faraj, S. and Huysman, M. (2021) 'Make Way for the Algorithms: Symbolic Actions and Change in a Regime of Knowing', Organization Science, 32(1), pp. 18–41.

Parker, G.G., Van Alstyne, M.W. and Choudary, S.P. (2016) Platform Revolution: How Networked Markets Are Transforming the Economy—and How to Make Them Work for You. New York: W.W. Norton & Company.

Pasquale, F. (2015) The Black Box Society: The Secret Algorithms That Control Money and Information. Cambridge, MA: Harvard University Press.

Polanyi, M. (1966) The Tacit Dimension. New York: Doubleday.

Shannon, C.E. (1948) 'A Mathematical Theory of Communication', Bell System Technical Journal, 27(3), pp. 379–423.

Sorensen, T., Moore, J., Fisher, J., Gordon, M.L., Mireshghallah, N., Rytting, C.M., Ye, A., Jiang, L., Lu, X., Dziri, N., Althoff, T. and Choi, Y. (2024) 'Position: A Roadmap to Pluralistic Alignment', in Proceedings of the 41st International Conference on Machine Learning. Proceedings of Machine Learning Research, 235, pp. 46280–46302.

Susskind, R. and Susskind, D. (2015) The Future of the Professions: How Technology Will Transform the Work of Human Experts. Oxford: Oxford University Press.

Varoufakis, Y. (2023) Technofeudalism: What Killed Capitalism. London: Bodley Head.

Weller, O., Boratko, M., Naim, I. and Lee, J. (2025) 'On the Theoretical Limitations of Embedding-Based Retrieval', arXiv preprint, arXiv:2508.21038.

Organisational AI Literacy and Readiness: Enabling Factors and Barriers

Riccardo Michele Colangelo, Eleonora Petrazzuoli, Maria Menshikova
Universitas Mercatorum, Rome, Italy

Abstract

The spread of artificial intelligence (AI) within organisations is reshaping how they operate, leading to deep and systemic transformations across processes, structures and decision-making practices.
In this context, the distinct concepts of AI literacy and AI readiness are increasingly central to understanding organisational AI-driven transformation.
This study investigates the relationship between AI literacy and AI readiness within organisations, also examining which enabling factors and barriers influence these constructs across multiple organisational levels.
By adopting an integrated bibliometric and systematic approach, this study aims to provide an overview of how these constructs are addressed in the literature, followed by a systematic analysis of enabling factors and barriers to AI literacy and AI readiness, with specific regard to 61 Q1 journal articles in the business, management, and accounting domain.
The bibliometric analysis, conducted using Biblioshiny, maps publication trends, geographical distribution and keyword co-occurrence networks.
The systematic analysis further explores enabling factors and barriers to AI literacy and AI readiness on a targeted selection of relevant peer-reviewed publications, using a specific purpose-built analytical framework, across micro (individual), meso (team), and macro (organisational) levels.
The key findings indicate that AI literacy is primarily associated with individual-level capabilities, including AI knowledge, self-efficacy and critical thinking skills, while AI readiness is predominantly linked to organisational dimensions such as technological infrastructure, data availability, leadership support and strategic alignment. Key barriers include skill gaps, cognitive biases, resource constraints, legacy systems, organisational rigidity, and regulatory uncertainty.
The study also identifies cross-cutting factors, such as trust in AI systems, ethical considerations and algorithmic bias, which influence both constructs and reflect the socio-technical nature of AI adoption.
Overall, the findings suggest that AI literacy and AI readiness are strongly interdependent constructs: they are closely interconnected and often mutually reinforcing in shaping how organisations engage with artificial intelligence, although they remain distinct constructs.

Keywords: AI literacy, AI readiness, Organisations, Enabling factors, Barriers

Paper type: Academic Research Paper

1 Introduction

The rapid diffusion of artificial intelligence (AI) across organisations is driving a profound AI-driven transformation that affects skills, competencies, decision-making processes, governance structures, and regulatory compliance. In this evolving landscape, AI literacy is increasingly framed not merely as an individual technical skill, but as a strategic organisational capability, not limited to decision-makers.

AI literacy, while remaining central to the research agenda (Jiang et al., 2024), lends itself to multiple definitions. The current state of the literature reveals a lack of a widely applicable definition of AI literacy (Liu et al., 2025), as highlighted by several systematic literature reviews documented in the scholarly literature (Lintner, 2024).

Indeed, AI literacy has been defined as “comprehensive proficiency” related to the “understanding, using, and working with AI” (Steinhauser and Heid, 2026), which “enables individuals to develop effective solutions” (Yordudom et. al, 2026) and “encompasses a crucial set of skills” (Roozen et al., 2025). Recent studies highlight that “AI literacy transcends fundamental digital literacy” (Lee et al., 2026), enabling “individuals to understand, critically assess, and purposefully engage with AI” (Testa et al., 2026).

The literature often combines or overlaps AI literacy with related constructs such as AI readiness (Jiang et al., 2024), which has also been examined in specific studies within the academic literature (Kaewsawad et al., 2025).

While, in the literature, readiness is conceptualized as “conditions that enable organizations to acquire, assimilate, and apply new knowledge” (Steinhauser and Heid, 2026), AI readiness concerns measures aimed to the preparation “for the changes involved in technology, infrastructure” (Sharma et al., 2022)” and it “can be understood as the extent to which individuals are prepared to engage with AI technologies in applied contexts”

(Testa et al., 2026). It should be acknowledged that "AI readiness is multidimensional" (Holmström, 2022) and that the readiness itself "may be either organizational or individual readiness" (Issa et al., 2022).

This study examines the relationship between AI literacy and AI readiness, while also identifying enabling factors and barriers emerging from the most relevant literature in business, management and accounting.

It contributes to the ongoing academic debate on these constructs by advancing the state of the art beyond the boundaries of a conventional systematic literature review.

The paper addresses the following research questions:

- *RQ1: There are boundaries and intersections between AI literacy and AI readiness? Do these notions represent distinct analytical dimensions or complementary facets?*
- *RQ2: What factors enable or hinder the development of AI literacy and AI readiness within organisations?*

The structure of the paper is as follows. Section 2 outlines the research methodology. Section 3 presents bibliometric evidence, both at a general level and within the specific thematic domain investigated. Section 4 provides a systematic content analysis of selected articles, based on a purposefully designed analytical framework aimed at identifying multi-level enabling factors and barriers associated with the two focal constructs. Section 5 discusses the findings, with particular emphasis on insights derived from the bibliometric analysis. Finally, Section 6 concludes the paper by summarising the main contributions and outlining avenues for future research.

2 Methodology

This paper aims to address the research questions through a combination of bibliometric and systematic analysis of selected documents within the business, management and accounting domain.

The present study draws on documents indexed in the Scopus database. The dataset was selected following the PRISMA framework, as illustrated in Figure 1.

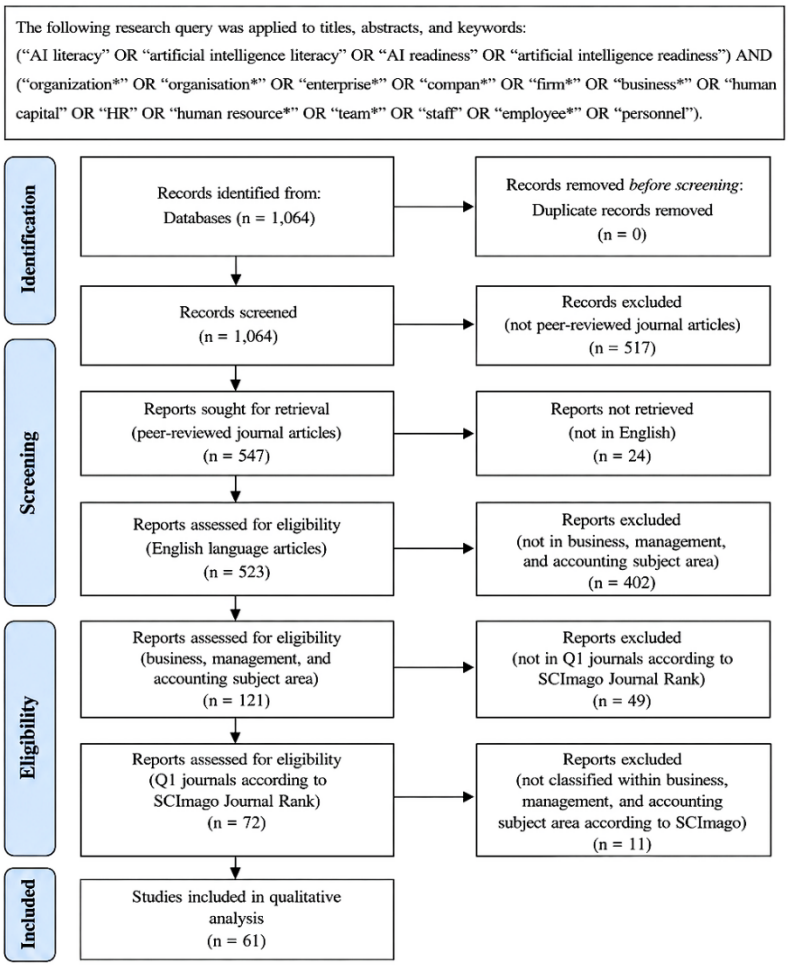


Figure 1: PRISMA framework applied to the reference dataset

The following research query was applied to titles, abstracts, and keywords:

("AI literacy" OR "artificial intelligence literacy" OR "AI readiness" OR "artificial intelligence readiness") AND ("organization" OR "organisation*" OR "enterprise*" OR "compan*" OR "firm*" OR "business*" OR "human capital" OR "HR" OR "human resource*" OR "team*" OR "staff" OR "employee*" OR "personnel").*

The initial search yielded 1,064 documents. The same number of contributions was obtained when considering the time span from 2020 to 2026. Restricting the sample to peer reviewed journal articles resulted in 547 documents, of which 523 were in English. This subset was used for the preliminary bibliometric analysis.

After further filtering, aimed to ensure domain specificity, the analysis focused on the business, management and accounting subject area, resulting in a final subset of 121 documents. This subset was used for a second, domain-specific bibliometric analysis.

In order to ensure the inclusion of relevant contributions, within the specific CSV subset, only articles published in Q1-ranked journals were selected using the SCImago Journal Rank, resulting in 72 documents. After excluding 11 articles published in Q1 journals that were not classified within the business, management, and accounting subject area according to SCImago, the final sample was reduced to 61 contributions. This final set of documents was subsequently used for the systematic analysis.

For the bibliometric analysis, Biblioshiny is employed in order to identify publication trends over time, country-level scientific production and the co-occurrence network of authors' keywords. This approach enables both a general overview of the intellectual and conceptual development and a specific insight regarding the abovementioned field.

The systematic analysis, in contrast, focuses on identifying the enabling factors and barriers related to both AI literacy and AI readiness within the specific context of business, management, and accounting. In particular, this analysis examines these factors across three levels: micro (individual), meso (team), and macro (organizational). This multi-level perspective is adopted to highlight similarities and differences in enabling conditions and barriers across the two theoretical constructs.

3 Bibliometric Analysis

Prior to the systematic analysis related to enabling factors and barriers to AI literacy and AI readiness, the bibliometric approach provides a broad-level overview and subsequently enables a more fine-grained focus on the business, management and accounting domain.

In particular, it allows for the identification of the evolution of scientific production over time, the mapping of its geographical distribution through the identification of the most productive countries and the generation of meaningful insights.

Moreover, co-occurrence mapping based on authors' keywords helps to uncover the intellectual structure of the field and the main thematic connections within the literature.

3.1 General bibliometric analysis

Prior to analysing articles specifically related to the field of business management and accounting, it is appropriate to carry out a preliminary and broader bibliometric analysis.

This analysis is based on a dataset of 523 articles retrieved from Scopus , covering the most recent decade (2016–2026). More specifically, these contributions represent a subset of the initial dataset and, without considering specific subject areas, meet the following inclusion criteria: only journal articles (547) and, subsequently, only publications in English (523).

The results of the analysis are as follows.

The bibliometric analysis shows a growing annual distribution of research articles on AI literacy and AI readiness in organisations, as illustrated in Figure 2.

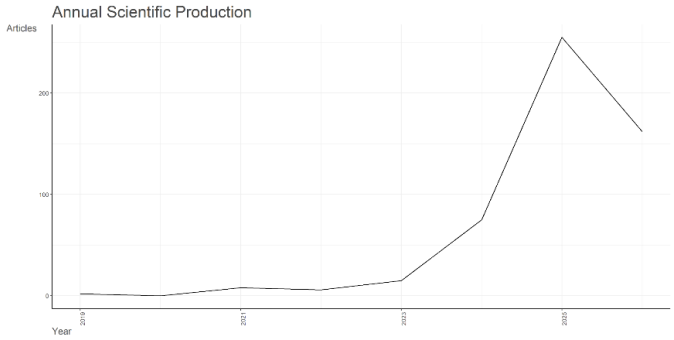


Figure 2: Trend of publication

Table 1 reports the yearly publication output on AI literacy and AI readiness in organisations over the analysed period (2016–2026). The data show no publications between 2016 and 2018, followed by the first emerging contributions in 2019 (2 articles). After a brief interruption in 2020 (0 articles), an increase can be observed in 2021 with 8 publications, followed by a slight decline in 2022 (6 articles). From 2023 onwards, the field experiences a clear acceleration, with 15 articles published in 2023 and a substantial jump in 2024 (75 articles). This growth culminates in 2025, with 255 publications. Although 2026 data are still partial, the 162 articles already recorded confirm the continuation of this strong upward trajectory, suggesting intensifying academic attention in this domain.

Table 1: Number of articles per year

Year	Articles
2016	0
2017	0
2018	0
2019	2
2020	0
2021	8
2022	6
2023	15
2024	75
2025	255
2026	162

Figure 3 illustrates a Country-wise production map, in which colour intensity reflects the number of articles, with darker tones representing higher publication volumes at the national level.

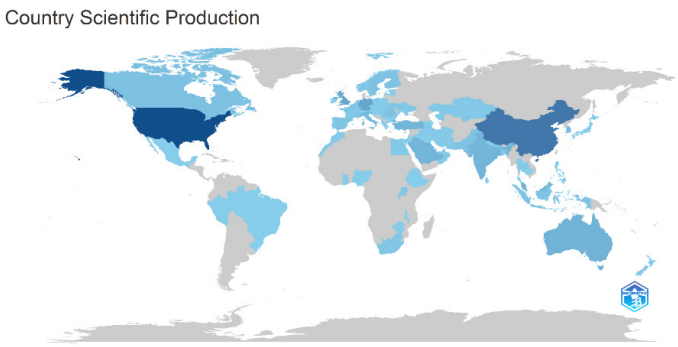


Figure 3: Country-wise production map

Table 2 highlights the top 10 countries that have made the most significant contributions to the topic. The data reveal a strong concentration of scientific production in the United States (357 articles) and China (235 articles), followed by the United Kingdom (108) and Germany (105). Other relevant contributors include Australia (72), Saudi Arabia (67), Turkey (67), India (66), Malaysia (64), and Italy (50).

Table 2: Top 10 countries contributing maximum number of articles

Country	Number of articles
USA	357
CHINA	235
UK	108
GERMANY	105
AUSTRALIA	72
SAUDI ARABIA	67
TURKEY	67
INDIA	66
MALAYSIA	64
ITALY	50

Bibliometric analysis also includes co-occurrence analysis of authors' keywords, enabling the identification of dominant thematic clusters and the relationships among emerging research topics (Agbo et al., 2021) as illustrated in Figure 4.

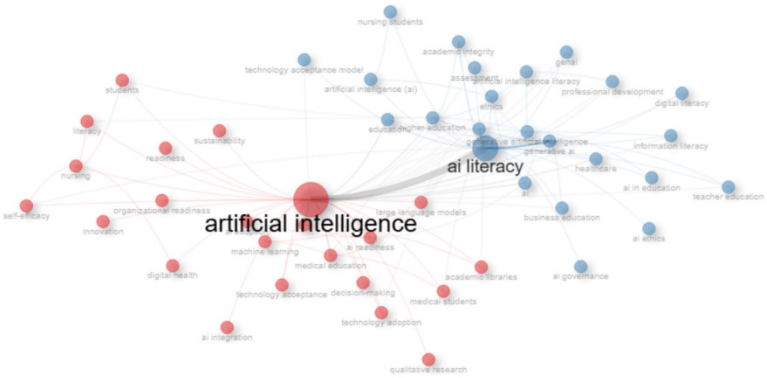


Figure 4: Keywords network (authors' keywords)

The co-occurrence network of authors' keywords reveals a structured thematic landscape organized around two main clusters

The red cluster is centered on artificial intelligence, which emerges as the dominant and most connected node in the network. This cluster primarily reflects a perspective centered on AI adoption and application, also in specific fields (e.g. healthcare). This cluster contains terms such as technology adoption, organizational readiness, AI readiness, and self-efficacy.

In contrast, the blue cluster is organized around AI literacy, which appears as a highly prominent concept. This cluster reflects a focus on the skills and knowledge required to effectively and responsibly engage with AI technologies and contains key terms such as artificial intelligence literacy and generative AI. At the same time, the inclusion of AI ethics and AI governance highlights increasing attention to the societal and normative implications of AI use. The educational dimension is further emphasized by keywords such as higher education, teacher education and professional development.

There are bridging keywords connecting the two clusters and indicating a progressive convergence between AI application and AI literacy discourses. Notably, the term literacy also appears in the red cluster, suggesting that competence-related dimensions are linked with AI readiness and adoption.

Overall, the increasing prominence of AI literacy suggests a shift from operational use toward critical engagement, competency development and responsible use and governance of AI.

3.2 Bibliometric analysis in the field of business, management and accounting

After this preliminary bibliometric analysis, it is appropriate to further examine the same bibliometric dimensions on a more restricted sample, comprising only articles related to the field of business, management and accounting. This inclusion criterion was specified in order to export a more focused and domain-specific database through Scopus.

This subset consists of 121 articles and enables the identification of specific bibliometric trends within the field of analysis.

The annual distribution of research articles on AI literacy and AI readiness in organisations is illustrated in Figure 3 and it is highly consistent with patterns observed in Figure 2, regardless of the specific field.

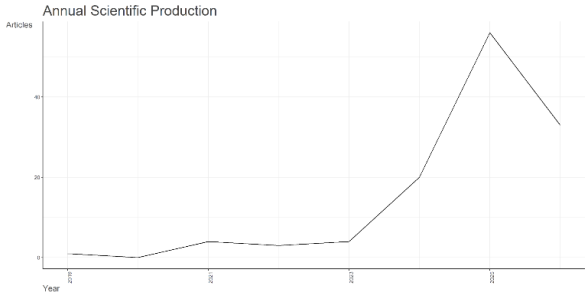


Figure 5: Trend of publication

Table 3 reports the corresponding yearly publication output within the specific field over the period analysed. The results indicate a gradual increase starting in 2021. A marked growth is observed from 2024 onwards, with a substantial peak in 2025, suggesting a rapidly accelerating research interest in the topic. Although 2026 data are partial, they already confirm the continuation of this upward trend.

Table 3: Number of articles per year

Year	Articles
2016	0
2017	0
2018	0
2019	1
2020	0
2021	4
2022	3
2023	4
2024	20
2025	56
2026	33

Figure 6 provides a visual representation of the most productive countries.

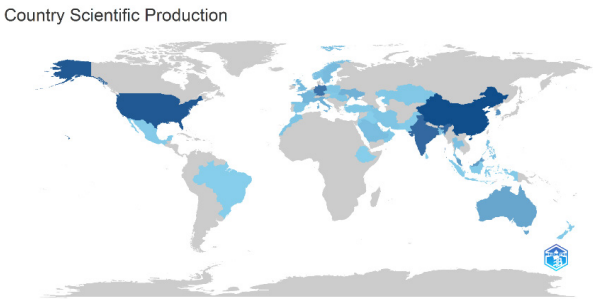


Figure 6: Country-wise production map

Table 4 highlights the top 10 countries that have made the most significant contributions to the topic within the specific field. The analysis reveals that China holds the leading position, contributing a total of 45 articles, followed closely by the United States with 41 publications and India with 35. Germany ranks fourth with 31 articles, while South Korea contributes 21 publications. Malaysia accounts for 17 articles, followed by Australia with 14 and Italy with 11. Ukraine and France complete the top ten, with 10 and 9 publications respectively. Overall, the results indicate a strong geographical concentration of research activity in Asia, North America, and selected European countries.

Table 4: Top 10 countries contributing maximum number of articles

Country	Number of articles
CHINA	45
USA	41
INDIA	35
GERMANY	31
SOUTH KOREA	21
MALAYSIA	17
AUSTRALIA	14
ITALY	11
UKRAINE	10
FRANCE	9

The bibliometric analysis is complemented by a co-occurrence analysis of authors' keywords, shown in Figure7.

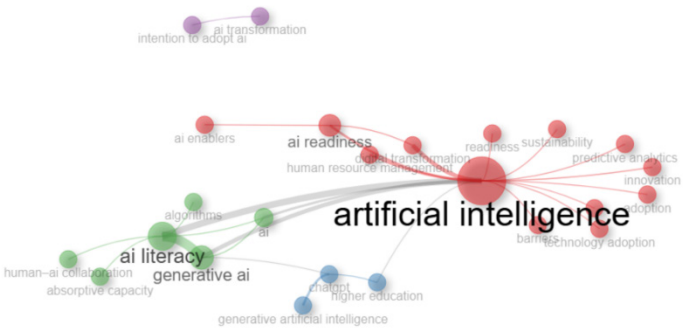


Figure 7: Keywords network (authors' keywords)

This keyword co-occurrence map reveals 4 distinct clusters, each reflecting a thematic area.

The red cluster is dominant and centered on artificial intelligence. It captures strategic and organisational dimensions of AI implementation. Key themes include, for example, AI readiness, digital transformation, technology adoption, innovation, predictive analytics and sustainability. It also reflects barriers and enablers at the organisational level.

The green cluster revolves around AI literacy and generative AI, including topics like algorithms, human–AI collaboration, and absorptive capacity, suggesting a focus on knowledge acquisition and user competences.

The blue cluster is smaller and more application-oriented, suggesting a link between higher education and AI literacy.

The purple cluster focuses on the early-stage perspective of AI integration, highlighting intention to adopt AI.

Overall, the network suggests a thematic progression: from intent and transformation (purple), organisational AI readiness and adoption (red), AI literacy development (green) and practical applications, especially in education (blue).

4 Systematic Analysis

As outlined in the methodology section, following the bibliometric analysis, the subset of 121 articles related to the field of business, management and accounting was identified in order to conduct a more in-depth systematic analysis.

For this purpose, journal quality was assessed using the SCImago Journal Rank, retaining only Q1-ranked journals from the specific CSV subset (121), which resulted in 72 articles. After excluding 11 articles published in Q1 journals not classified within the business, management and accounting subject area according to SCImago, the final sample was reduced to 61 contributions.

The systematic analysis was ultimately conducted on 61 articles.

These articles were analysed through a specifically purpose-built analytical framework designed to identify enabling factors and barriers at the micro (e.g. employees), meso (e.g. teams), and macro (e.g. organisation) levels, with reference to AI literacy or AI readiness within the specific context outlined above.

In particular, studies related to AI literacy are predominantly concentrated at the micro level, focusing on individual competencies, cognitive capabilities, and human–AI interaction. In this domain, the most recurrent enabling factors include self-efficacy (Duong, 2026), AI knowledge, critical evaluation skills (Kulkarni at al, 2024), and usage capabilities, while the main barriers concern low levels of competence, difficulties in understanding AI systems, cognitive biases, and over-reliance on automated outputs (Jeremiah, 2025).

In contrast, AI readiness primarily emerges at the meso and macro levels, with a focus on the organisational and systemic capabilities required for the adoption and implementation of AI. In these studies, enabling factors include technological infrastructure (Nguyen at al., 2024), data availability and quality, top management support, strategic alignment, and organizational capabilities, whereas the most relevant barriers are represented by resource constraints, high costs, legacy systems, lack of specialized skills, organizational rigidity, and regulatory uncertainty (Kumar at al., 2021).

Two tables presenting the main enabling factors and barriers related to AI literacy and AI readiness are provided in the paper's appendix.

5 Discussion

The analysis also highlights that some contributions adopt a multi-level perspective, integrating both individual and organizational dimensions. In such cases, AI literacy is frequently identified as an enabling factor for AI readiness, suggesting a hierarchical or sequential relationship between the two constructs (Baabdullah at al., 2021). In particular, individual competencies contribute to strengthening the organizational capacity to effectively adopt and leverage AI technologies.

Finally, increasing attention is devoted to cross-cutting factors, such as trust in AI systems, ethical considerations, algorithmic bias, and transformations in decision-making processes, which influence both literacy and readiness, highlighting the socio-technical nature of artificial intelligence adoption. Trust, for instance, is frequently identified as a key mediator at the individual level, shaping users' willingness to engage with AI systems, while at the organizational level it becomes intertwined with governance, transparency, and accountability mechanisms (Vafaei-Zadeh at al., 2025). Similarly, ethical concerns, ranging from bias and fairness to data privacy, are increasingly framed not only as risks or barriers, but also as conditions for sustainable and responsible AI deployment.

Finally, the literature underscores that the adoption of AI entails a reconfiguration of decision-making processes, where human judgment and algorithmic outputs become progressively intertwined. This transformation is particularly evident in studies on organizational AI readiness and digital transformation, which show how AI reshapes organizational activities, structures, and strategic goals (Holmström, 2022). At the same time, research on human-AI interaction highlights the emergence of hybrid forms of decision-making, in which AI systems actively collaborate with human actors, challenging traditional notions of autonomy, expertise, and control.

These findings suggest that AI adoption cannot be fully understood through a purely technological lens, but should instead be interpreted as a complex socio-technical phenomenon, in which individual competencies and organizational capabilities co-evolve. In particular, studies explicitly linking AI literacy and AI readiness highlight how individual-level capabilities, such as AI knowledge, self-efficacy, and critical understanding, contribute to shaping organizational readiness and implementation capacity (Issa at al., 2022).

6 Conclusion

This study explored the relationships between AI literacy and AI readiness, with particular attention to their respective enabling factors and barriers. It also employed a bibliometric approach alongside a more in-depth systematic analysis of Q1 journal articles in the fields of business management and accounting.

The analysis focused on enabling factors and barriers related to three different levels within organizations and led to several key findings in response to the research questions posed.

RQ1: AI literacy and AI readiness are distinct yet strongly interdependent constructs, operating across different levels of analysis but converging within a common socio-technical framework. While AI literacy primarily concerns individual-level capabilities related to understanding and interacting with AI, AI readiness reflects the organizational capacity to adopt and implement these technologies.

RQ2: The development of AI literacy and AI readiness is influenced by a multi-level set of enabling factors and barriers, reflecting the socio-technical nature of AI adoption. At the individual level, capabilities such as AI knowledge, self-efficacy, and critical evaluation skills emerge as key enablers, while skill gaps, cognitive biases, and limited understanding represent major obstacles. At the organizational level, factors such as technological infrastructure, data availability, leadership support, and strategic alignment are critical in enabling AI readiness,

whereas resource constraints, legacy systems, organizational rigidity, and regulatory uncertainty act as significant barriers. Additionally, cross-cutting elements, such as trust, ethical considerations, and algorithmic bias, play a pivotal role across both dimensions. Overall, the findings suggest that successful AI adoption depends on the alignment of individual competencies and organizational capabilities, supported by appropriate governance and ethical frameworks.

This study is characterised by several limitations. In particular, the present analysis is exclusively based on the Scopus database, which, despite its broad coverage and scholarly reliability, may nevertheless impose certain constraints on a more comprehensive systematic exploration of the investigated themes. Indeed, relevant studies and contributions might exist outside the Scopus-indexed corpus and therefore remain unconsidered in this review.

The analyses conducted in this article aim to provide research insights that may contribute to enriching the scientific debate on AI literacy and AI readiness, conceptualized as two distinct yet interrelated constructs.

The identification of shared enabling factors and barriers offers a basis for future research, which should further explore the reciprocal relationships between these constructs, moving beyond the boundaries of each individual domain. Future research may also address the inherent limitations of this study by expanding the analytical dataset.

Furthermore, the exclusion of book chapters and other document types, as well as contributions published in non-Q1 journals and not related to the fields of business management and accounting, may have resulted in the omission of potentially significant contributions to the topic under investigation.

Furthermore, future research could extend both the scope of the search strategy and the breadth of the research area by incorporating additional relevant keywords. It may also be beneficial to include other kinds of contributions, which could offer valuable and complementary insights to the existing body of literature.

7 Ethics declaration

No ethics declaration needed.

8 AI declaration

No AI tools were used in the development of this paper.

References

Agbo, F.J., Oyelere, S.S., Suhonen, J. and Tukiainen, M. (2021) "Scientific production and thematic breakthroughs in smart learning environments: a bibliometric analysis", *Smart Learning Environments*, Vol. 8, No. 1, pp 1-25

Baabdullah, A. M., Alalwan, A. A., Slade, E. L., Raman, R. and Khatatneh, K. F. (2021) "SMEs and artificial intelligence (AI): Antecedents and consequences of AI-based B2B practices", *Industrial Marketing Management*, Vol. 98, pp 255-270

Duong, C.D. (2026) "How do Artificial Intelligence-facilitated enablers influence international entrepreneurship: Does the Trump 2.0 tariff policy matter?", *Entrepreneurial Business and Economics Review*, Vol. 14, No. 1, pp 177-194

Holmström, J. (2022) "From AI to digital transformation: The AI readiness framework", *Business Horizons*, Vol. 65, No. 3, pp 329-339

Issa, H., Jabbouri, R. and Palmer M. (2022) "An artificial intelligence (AI)-readiness and adoption framework for AgriTech firms", *Technological Forecasting and Social Change*, Vol. 182, 121874

Jeremiah, F. (2025), "The human-AI dyad: Navigating the new frontier of entrepreneurial discourse", BENCHMARKING, Vol. 166, 103529

Jiang, T., Sun, Z., Fu, S., and Lv, Y. (2024) "Human–AI interaction research agenda: A user-centered perspective", *Data and Information Management*, Vol. 8, No. 4, 100078

Kaewsawad, V. and Duda, J. (2025) "Assessing artificial intelligence (AI) literacy and readiness in Thailand's workforce: Challenges and opportunities for digital transformation", *Science and Technology Asia*, Vol. 30, No. 3, pp 209-231

Kulkarni, A.V., Joseph, S. and Patil, K.P. (2024), "Artificial intelligence technology readiness for social sustainability and business ethics: Evidence from MSMEs in developing nations", *International Journal of Information Management Data Insights*, Vol. 4, No. 2, 100250

Kumar, S., Raut, R D., Queiroz, M. M. and Narkhede, B. E. (2021) "Mapping the barriers of AI implementations in the public distribution system: The Indian experience", *Technology in Society*, Vol. 67, 101737

Lee, S.C., Baby, T., Vongvit, R., Lee, J., Kim, Y.W., Cha, M.C. and Yoon, S.H. (2026) "Development and validation of Generative AI Competence Scale (GenAIComp) among university students", *Technology in Society*, Vol. 84, 103059

Lintner, T. (2024) "A systematic review of AI literacy scales", *npj Science of Learning*, Vol. 9, No. 1, 50

Liu, X., Zhang, L., & Wei, X. (2025) "Generative artificial intelligence literacy: Scale development and its effect on job performance", *Behavioral Sciences*, Vol. 15, No. 6, 811

Nguyen, T. T., Dao, T. T., Tran, T. B., Nguyen, H. T. T. Le, L. T. N. and Pham, N. T. T. (2024) "Fintech literacy and digital entrepreneurial intention: Mediator and Moderator Effect", *International Journal of Information Management Data Insights*, Vol. 4, No. 1, 100222

Roozen, I. Raedts, M., Claeys, C. and Di Gennaro, G. (2025), "Gendered AI in banking services: the influence of financial chatbots' gender on consumer behaviour", International Journal of Bank Marketing, Vol. 43, No. 6, pp 1231-1251

Sharma, K., Jain, M. and Dhir, S. (2022), "Analysing the impact of artificial intelligence on the competitiveness of tourism firms: a modified total interpretive structural modeling (m-TISM) approach", *International Journal of Emerging Markets*, Vol. 17, No. 4, pp 1067-1084

Steinhauser, S. and Heid, P. (2026) "Organizational Readiness, AI Literacy, and the New Frontier of R&D: How Generative AI Shapes Innovation Capacity", *R and D Management*

Testa, M.; Apuzzo, A. and Pittaway, L. (2026) AI literacy alone is not enough: Student AI readiness and career adaptability in business and management education, *International Journal of Management Education*, Vol. 24, No. 2, 101394

Vafaei-Zadeh, A., Nikbin D., Wong S.L. and Hanifah H. (2025) "Investigating factors influencing AI customer service adoption: an integrated model of stimulus–organism–response (SOR) and task-technology fit (TTF) theory", *Asia Pacific Journal of Marketing and Logistics*, Vol. 37, No. 6, pp 1465-1502

Yordudom, T, Nui-Suk, C., Imjai, N., Sumiati, S. and Aujirapongpan, S. (2026) "Developing open innovation competency in Gen Z entrepreneurship students: Roles of absorptive capacity, AI literacy and algorithmic thinking skills", *International Journal of Management Education*, Vol. 24, No. 1, 101326

Annex 1: Enabling factors and barriers to AI literacy

AI Literacy		
Level	**Enabling actors**	**Barriers**
Micro	Fintech literacy dimensions, technical capacity; critical thinking; professional experience; AI-driven technologies; communication and collaboration; safety and ethics; problem-solving; AI self-efficacy; AI competency; AI knowledge; perceived desirability; Perceived feasibility; attitude towards entrepreneurship; global mindset; AI understanding; human–AI communication capability and interaction; human-AI collaboration; data-driven insights; individual competencies; cognitive capabilities.	Lower tech infrastructure and culture; data privacy; lack of transparency; information asymmetry; strong dependence on AI; tariff policy uncertainty; environmental uncertainty; perceived risk; cognitive dissonance (desirability–feasibility mismatch); ethical concerns; algorithmic bias; identity tension; feedback loop bias.
Meso	Algorithmic literacy; decentralization of algorithmic skills; workforce training/reskilling; AI infrastructure; Generative AI e LLM; AI competencies; ethical AI frameworks; AI value identification; AI implementation ability; HR-related AI capability; executive expertise; decision scrutiny; managerial attention; resource allocation; executive power; technological readiness; cross-functional collaboration; leadership support; data integration.	Digital gap; bias; over-dependency AI; lack of AI orientation; limited AI implementation ability; minimal AI business impact; skill gap; rigid business models; resistance to change; lack of integration; low knowledge; AI complexity.
Macro		

Annex 2: Enabling factors and barriers to AI readiness

AI Readiness		
Level	**Enabling actors**	**Barriers**
Micro	Social Influence; emotional trust; task-technology Fit.	Psychological and motivational resources; generational gap.
Meso	Technological and organizational capability; critically interpret algorithmic outputs; value-oriented capabilities; governance; understanding AI limitations; Interactivity; autonomy; mobility; human-AI collaboration; data capabilities; technology roadmapping; infrastructure; attitude; awareness; professional expertise.	Gap between technological adoption and value realization; rely uncritically on algorithmic outputs; multiple stakeholders; lack complete human-level language abilities; organizational barriers; strategic and economic barriers; AI phobia; AI fever; black box AI; governance; high cost; legacy systems; need for appropriate skill; social order
Macro	Scientific research output; internet infrastructure; public consumption expense; education; governance; gender equality; organizational readiness; structural readiness; cultural readiness; digital infrastructure; internet connection; sensor networks; institutional coordination; public-private partnerships; machine learning and analytics; digital capability.	Technical barriers; limitations in using generative AI tools; political issues; lack of economic policy and regulations; lack of IT infrastructure; limited digital infrastructure; connectivity barriers; low levels of digital literacy.

New Frontiers in Organizational Aesthetics

Giovanna Afeltra[1], Maria Menshikova[2], Maria Pia Ciano[3]
[1] University San Raffaele, Rome, Italy
[2] Mercatorum University, Rome, Italy
[3] Notthingham University, UK

Abstract

Over the past decades, Organizational Aesthetics (OA) has evolved from a marginal research stream into a vibrant interdisciplinary field addressing the sensory, effective, and material dimensions of organizing. Despite this growing interest, the literature remains fragmented across disciplines, methodologies, and levels of analysis, leaving key theoretical and empirical questions unresolved. This study responds to these challenges by offering a comprehensive Bibliometric-Systematic Literature Review (B-SLR) of OA research to map its intellectual structure, methodological trends, and emerging frontiers.

Drawing on a dataset of 82 peer-reviewed journal articles indexed in Scopus, the study combines bibliometric techniques (using Bibliometrix) with an in-depth qualitative analysis of selected contributions from leading journals. The findings show a steady growth of OA research since the mid-1990s, with strong contributions from the United Kingdom and the United States, alongside increasing global diversification. Thematic analysis highlights core areas such as embodiment, emotions, and aesthetic labour, while identifying emerging domains including workplace aesthetics, employee well-being, and environmental design.

The review reveals that OA has shifted from a focus on beauty and artefacts toward a broader understanding of aesthetics as a constitutive, embodied, and political dimension of organizational life. Methodologically, the field remains dominated by qualitative and interpretive approaches, though recent studies signal a move toward quantitative and mixed-methods approaches. Importantly, the analysis uncovers a growing "aesthetic turn" in digital and hybrid work environments, where aesthetic experience extends beyond physical spaces into virtual and platform-mediated contexts.

Building on these insights, the paper proposes a future research agenda centered on theoretical integration, ethical and responsible organizing, methodological advancement, digital aesthetics, and cross-cultural perspectives. Overall, the study positions OA as a critical lens for understanding contemporary organisations and highlights its relevance for both scholars and practitioners seeking to design more human-centred, inclusive, and sustainable work environments.

Keywords: Organizational Aesthetics, Workplace Aesthetics, Work Environment Aesthetics, Office Aesthetics

Paper Type: Academic Research Paper

1 Introduction

Over the last decades, research on Organizational Aesthetics (OA) has moved from a niche concern to a interdisciplinary conversation about the sensory, affective, and material dimensions of organising shape strategy, coordination, identity, and power. Early studies showed that aesthetic cues - sounds, images, spatial design, dress and bodies, artefacts, and atmospheres - are not epiphenomena but constitutive of how organisations work in practice (Witz et al. 2003; Barry & Rerup, 2006; Bathurst, 2009). Subsequent contributions deepened this insight by foregrounding paradox, politics, and embodiment: aesthetics can humanise and enliven organising, but it also frames and disciplines conduct through tastes, styles, and senses (Linstead, 2018; Strati, 2019; Strati, 2025).

In parallel, scholars have examined aesthetic labour and gendered embodiment in service and cultural industries (Witz et al., 2003; Warren, 2008), cinematic and visual framings of organising (Barry & Rerup, 2006), music and emotion in work (Richard & Reveley, 2007), and the role of office quality, ambience, and atmospheres for performance and well-being (Teven & Comadena, 1996). More recently, OA has intersected with information systems, urban and spatial studies, and contemporary workplace transformations, remote/hybrid arrangements, co-working infrastructures, and the aesthetic governance of digital platforms (Du et al. 2024; Ricciardelli et al. 2024; Tuzcuoğlu et al. 2025). These developments signal a field that is simultaneously consolidating core constructs and pushing into new terrains related to temporary projects, power and political aesthetics, as well as the moral ambivalence of taste and style (Raadik & Kuura, 2025; Strati, 2025; Chang et al., 2023).

Despite this expansion, literature remains fragmented across disciplines (organisation studies, sociology, philosophy of management, information systems, communication, urban studies), approaches (ethnography, visual analysis, discourse analysis, experiments, surveys), and levels of analysis (micro-embodiment, meso-

spaces and artefacts, macro-cultural fields. Fundamental debates persist around: (1) whether aesthetics is primarily epistemic, an alternative way of knowing organizations, or ontological, constitutive of organizations themselves (Warren, 2008; Linstead, 2018); (2) how to reconcile the edifying and attractive promise of aesthetics with its more disciplinary and political functions (Strati, 2019, 2025); and (3) which methods best capture fleeting, atmospheric, and multisensory phenomena without reducing their essence (Warren, 2008; Barry & Rerup, 2006). New empirical sites, hybrid workscapes, digitally mediated collaboration, and the aestheticisation of everyday organisational life, have intensified these questions (Du et al., 2024; Tuzcuoğlu et al., 2025), while adjacent conversations on "aesthetic intelligence" in management, brand atmospherics, and religious or moralised forms of aesthetic labour widen the scope (Bjerke et al. 2010; Lewis & Aune, 2023).

This paper responds to these dynamics by offering a Bibliometric-Systematic Literature analysis of the OA corpus to map its evolution, intellectual structure, and emergent frontiers.

2 Methodology

This study adopts the Bibliometric-Systematic Literature Review (B-SLR) approach by Marzi et al. (2025), integrating the quantitative rigour of bibliometric analysis with the qualitative depth of systematic review. Combining analytical, synthetic, and theory-building phases, the process follows Marzi et al.'s 10-step framework to ensure rigour, transparency, and replicability.

Step 1: Definition of Research Question and Boundaries of the Study

Despite growing interest across disciplines, the concept of OA remains theoretically fragmented and empirically limited. This exploration informed the formulation of the central research question:

How has the concept of organizational aesthetics been conceptualized and investigated within management and organization studies?

Which countries have most influenced the development of organizational aesthetics as a field of study?

What research methodologies have been used to study aesthetics in organizations?

How is organizational aesthetics evolving in digital and hybrid work environments?

What are the possible future research directions in the field of organizational aesthetics?

Step 2: Search Query Definition

Drawing on existing terminology and alternative expressions found in previous works, the final query was formulated as: *"organizational aesthetics" OR "work environment aesthetics" OR "workplace aesthetics" OR "office aesthetics".* The search was tested and refined through trial queries to balance recall and relevance.

Step 3: Database Selection

Scopus was selected as the primary data source (Marzi et al., 2025). The search was executed in April 2026, retrieving an initial dataset of 136 documents.

Step 4: Data Screening

Applying the inclusion criteria (peer-reviewed, English-language journal articles) led to a refined dataset of 89 documents. Only peer-reviewed journal articles written in English were considered. Moreover, in accordance with PRISMA and AMSTAR 2 protocols, the dataset was screened to remove duplicates and ineligible records.

Step 5: Data Cleaning and Export

All retained records were manually reviewed to verify alignment with the research focus on OA. The final set of 82 articles was exported from Scopus in CSV format for bibliometric processing.

Step 6: Bibliometric Approach

The bibliometric analysis was performed using Bibliometrix to enable descriptive and network-based analyses, including co-occurrence of keywords.

Step 7: Cluster Identification and Thematic Mapping

The co-occurrence network generated by Bibliometrix was used to identify thematic clusters. The clusters were visually examined through network, density, and overlay visualizations to capture both established and emerging topics.

Step 8: Sample Ordering and Selection for Systematic Review

To deepen the analysis, the review focused on contributions published in the three top-ranked journals (according to SCImago) that most frequently addressed organizational aesthetics: Organization Studies, Organization, and Management Learning. From these journals, 17 articles were selected.

Step 9: Systematic Literature Review

A qualitative content analysis was conducted on the selected papers, examining their objectives, theories, methods, and findings.

Step 10: Theory Development and Future Research Agenda

In the final stage, the insights from both bibliometric and systematic analyses were synthesized to formulate a Future Research Agenda.

3 Results

3.1 Bibliometric analysis

The analysis shows a steady rise in OA research from 1996 to 2025. After an exploratory phase until the mid-2000s, publications grew rapidly between 2005 and 2011, with peaks in 2013, 2022 and 2025 (Fig.1). This trend confirms the field's progressive maturation as a recognised area within management and organisation studies.

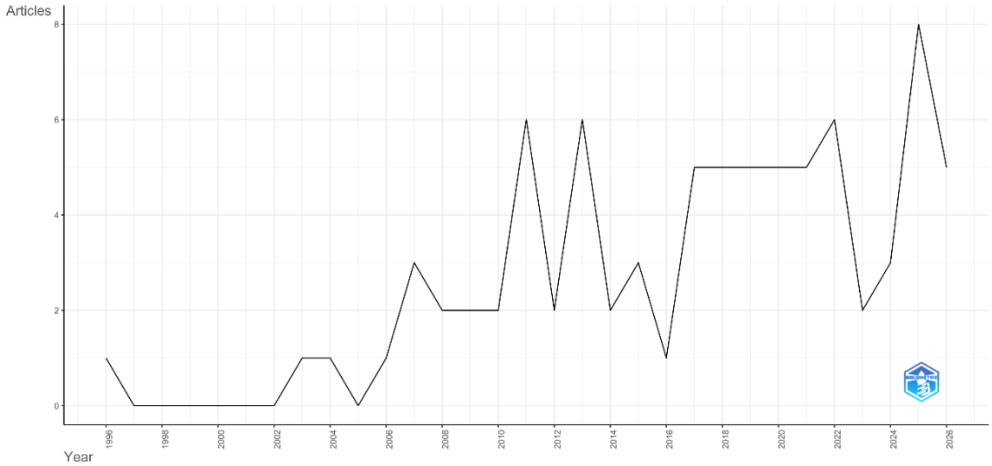


Figure 1. Annual Scientific Production
(Source: Authors' elaboration through Bibliometrix)

Regarding the geographical distribution of scientific production, the analysis identifies key contributors to OA research: the United Kingdom (21), United States (21), France (15), Brazil (10), and China (10), Israel (9), Netherland (9), Italy (8), Australia (7), Denmark (7). It is interesting to note that the studies in this field are present globally in European, Asiatic and American countries, enriching this domain with different cultural perseverance (Fig. 2).

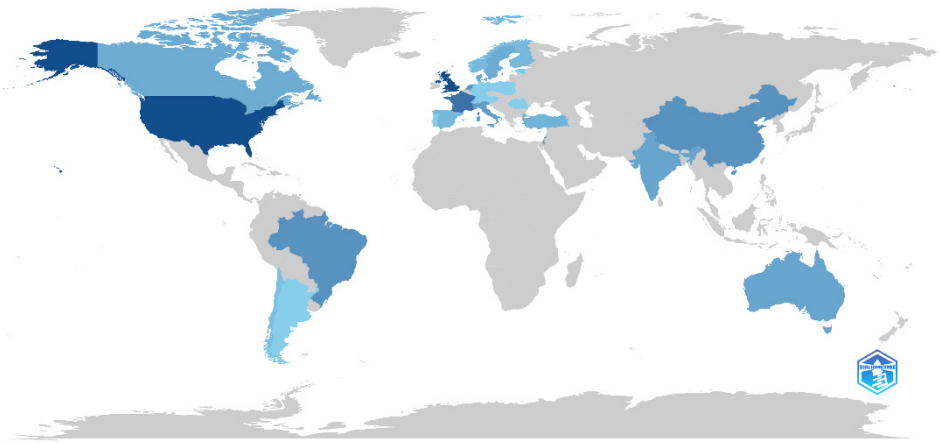


Figure 2. Country Scientific Production

Figure 3 highlights the main academic journals contributing to the literature on OA, showing a relatively concentrated distribution across a limited number of sources. In particular, "Organization Studies" emerges as the leading outlet, with the highest number of contributions (12 articles), confirming its central role for this research field. It is followed by "International Journal of Work Organisation" and "Emotion and Organization", each with three contributions. A second group of journals, including "Culture and Organization", "Journal of Management Inquiry", "Journal of Organizational Change Management", "Management Learning",

“Organization Science”, and “Studi di Estetica” each contributes two articles. “Academy of Management Annals” appears with a single contribution.

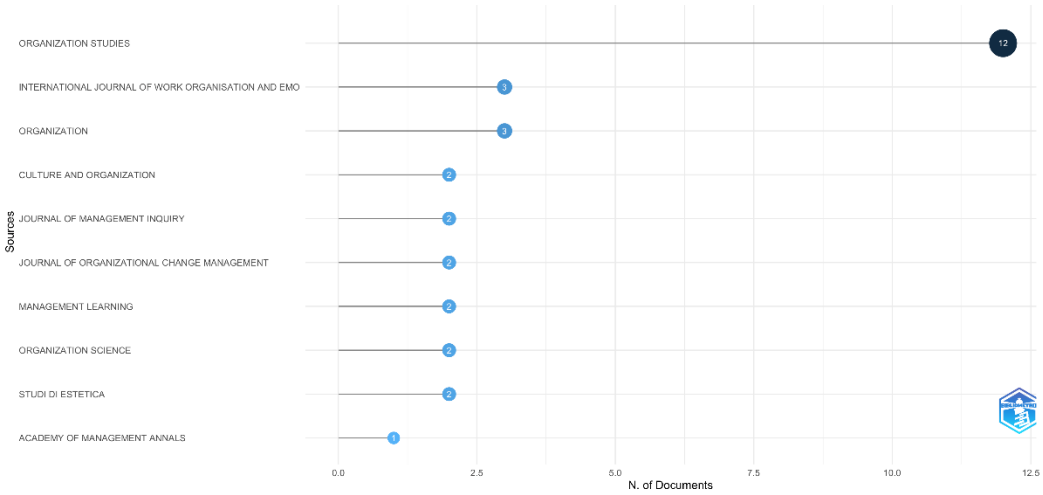


Figure 3. Most Relevant Sources
(Source: Authors’ elaboration through Bibliometrix)

Figure 4 presents a thematic map of literature, organizing key topics (authors’ keywords) according to their centrality (relevance within the field) and density (level of development). In the motor themes quadrant (high centrality and high density), topics such as organizational aesthetics, embodiment, and emotions emerge as well-developed and highly influential. These themes appear to play a pivotal role in shaping the intellectual structure of the field, suggesting a strong integration between aesthetic experience and embodied organizational processes. Closely related, the cluster including aesthetics, art, and aesthetic labour also occupies a central position, reinforcing the importance of artistic and expressive dimensions of organizational practices. The basic themes quadrant (high centrality but lower density) includes organizational aesthetics, creativity, and arts. These topics are foundational to the field and widely connected to other themes, yet they appear less internally developed, indicating opportunities for further theoretical consolidation. The niche themes quadrant (high density but lower centrality) is represented by topics such as employee well-being and office aesthetics. These areas are well developed but more specialized, suggesting focused streams of research that are not yet fully integrated into the broader discourse. The emerging or declining themes quadrant (low centrality and low density) includes environmental aesthetics, cultural and creative industries, and organizational design. These topics could represent both new research directions to be developed in the future and areas that are losing importance within the field.

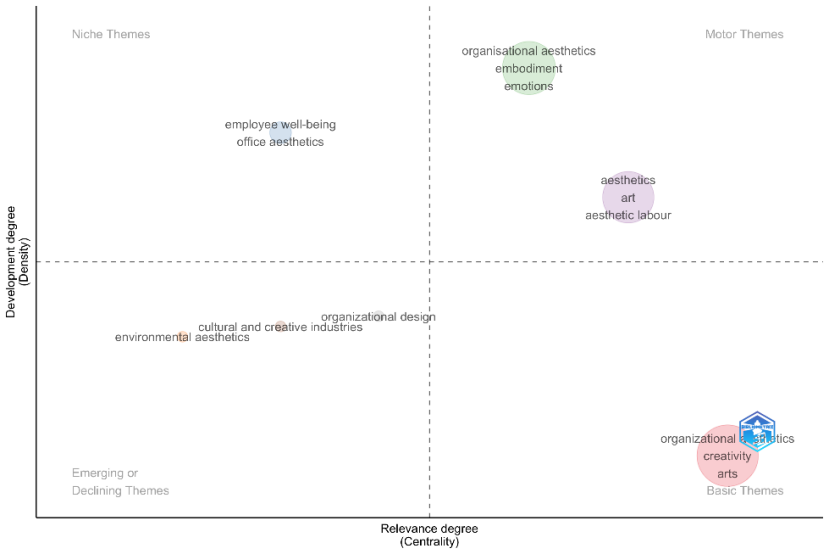


Figure 4. Thematic Map
(Source: Authors’ elaboration through Bibliometrix)

3.2 Systematic Analysis

3.2.1. Conceptualization and investigation of OA

Across the period covered by the dataset, OA has moved from a niche metaphor to a substantive analytic lens. Early and continuing contributions of Strati conceptualize aesthetics as embodied, sensory and experiential knowing in organizations, linking artefacts, spaces and everyday practices to how organizational members feel, judge and act (e.g., Strati 2019; Strati 2025). Building on this, research has treated aesthetics not simply as "beauty" but as a material–symbolic infrastructure through which organizations produce meaning.

A strong critical stream shows how aesthetics organizes identities and power. Several studies examine how spatial and visual regimes regulate inclusion/exclusion, identity work, and gendered displays (Wasserman & Frenkel 2011; 2015; Wasserman 2012; Sapir et al. 2025; Kachtan & Wasserman 2015). In parallel, classic work on aesthetic labour demonstrates how organizations script employees' looks and performances as part of value creation (Witz et al. 2003). Other authors connect OA to leadership, creativity and design logics (Barry & Rerup 2006; Bathurst 2009), or develop reflexive agendas for studying interactional and performative aesthetics (Warren 2008). Sustainability-oriented contributions argue for the value of beauty in responsible organizing (Ivanaj et al. 2018) and link urban/organizational design to social innovation (Ricciardelli et al. 2025).

3.2.2. Research methodologies used

Methodologically, the field is predominantly qualitative and interpretive. The dataset includes ethnography, interviews, visual/arts-based methods, and practice-based inquiry (e.g., Warren 2008; Strati 2019, Strati 2025). Critical-spatial analyses interrogate open-plan offices, transparency, and visibility as aesthetic–power mechanisms (Wasserman 2012; Wasserman & Frenkel 2015).

In recent years mixed and quantitative operationalizations have begun to appear, e.g. measurement of workplace aesthetic appreciation and links to strain/well-being (Nair et al. 2026). Field and quasi-experimental designs examine aesthetic cues in employment and HR contexts (Ronda & de Gracia 2022; Merriman et al. 2025). Built-environment/space studies use comparative workplace evaluations (Kirillova et al. 2020) and home-vs-office assessments (Tuzcuoğlu et al 2025).

3.2.3. OA in digital and hybrid work environments

Post-2020 documents in the dataset point to an aesthetic turn in hybrid and digital work. Some studies compare home and office aesthetics and their effects on satisfaction and productivity, highlighting how comfort, biophilic elements, and sensory cues travel (or not) across settings (Tuzcuoğlu et al. 2025; Kirillova et al. 2020). HR-facing work shows that aesthetic signaling in online job posts (e.g., biophilic imagery, visual tone) shapes applicant arousal and interest (Merriman et al. 2025), while office aesthetics influences job choice in competitive markets (Ronda & de Gracia 2022).

At the same time, some research underscores how visibility, exposure, and aesthetic control mutate in open/transparent workplaces, now also digital, with implications for identity regulation and gendered performances (Wasserman 2012; Wasserman & Frenkel 2011; Sapir et al. 2025; Vonk 2024). OA is currently being reconfigured through processes of digital mediation, in which aesthetic experience transcends material architecture and manifests in virtual interfaces, online platforms, and technologically framed embodiments.

3.2.4. Future research directions

Analysis of the documents in the dataset has identified several priorities for future research.

The first direction concerns theoretical integration and ethics. Scholars are encouraged to bridge OA with established frameworks such as sensemaking, institutional theory, and identity regulation to understand how aesthetic regimes contribute to the stabilization or disruption of organizational norms. Building on the work of Wasserman & Frenkel (2015), future research should also explore the intersection between aesthetics, ethics, and inclusion, examining how aesthetic arrangements influence issues of sustainability, fairness, and diversity within organizations (Ivanaj et al., 2018; Ricciardelli et al., 2025).

A second direction involves developing measurement tools and conducting causal analysis. While recent studies have begun to operationalize aesthetic experiences (e.g., Nair et al., 2026), further research is needed to design validated, multi-dimensional instruments that can capture environmental, interactional, and digital aesthetics. The use of field experiments and longitudinal designs could help identify the causal impact of aesthetic variables on creativity, learning, employee retention, and well-being.

A third emerging area relates to aesthetics in hybrid and digital environments. As organizations increasingly rely on virtual collaboration and hybrid work models, it becomes essential to investigate how digital platforms,

algorithmic interfaces, and camera-mediated interactions shape aesthetic experience and organizational sensemaking. Studies such as Tuzcuoğlu et al. (2025) and Merriman et al. (2025) highlight the need to explore how employees craft personal aesthetic ecologies that span across home, office, and virtual spaces.

Future research should also revisit questions of power, gender, and aesthetic labour in contemporary contexts. Classic insights on aesthetic labour (Witz et al. 2003) can be updated to address new organizational forms characterized by social media visibility, employer branding, and influencer-style communication.

Another promising avenue concerns cross-cultural and Global South perspectives. Research on OA is expanding geographically, comparative studies could uncover diverse cultural repertoires of taste, embodiment, and spatial sensibility, highlighting how these differences shape aesthetic practices and their organizational implications.

Finally, scholars should advance research on aesthetics for responsible and regenerative organising, linking aesthetics with sustainability and well-being. Building on Ivanaj et al. (2018) and Ricciardelli et al. (2024), future studies could explore how aesthetic design fosters restorative, biophilic, and ethical work environments, aligning beauty with ecological and social value creation.

6 Conclusions

This study provided a comprehensive overview of the intellectual evolution, thematic configuration, and methodological orientations of research on organisational aesthetics, offering both theoretical contributions and practical implications for management and organisational practice. By applying the Bibliometric-Systematic Literature Review (B-SLR) methodology (Marzi et al., 2025), the analysis combined quantitative mapping with qualitative synthesis to identify the main conceptual, methodological, and geographical trends that define the field.

The results show that the concept of organisational aesthetics has evolved from early interpretations centred on beauty, design, and artefacts toward a broader and more complex understanding of aesthetics as an embodied, affective, and political dimension of organising. In management and organisation studies, aesthetics is now conceptualised as a constitutive mechanism through which meaning, identity, and power are produced and experienced in organisations, rather than as a secondary or decorative element (Strati, 2019; Linstead, 2018; Warren, 2008).

From a geographical perspective, the field has been primarily shaped by scholarship from the United Kingdom and the United States, reflecting their strong traditions in critical and interpretive organisation theory. In recent years, significant contributions have also emerged from other European (e.g. Netherland, Italy) and non-European countries (Brazil, China, Israele) signalling a growing internationalisation and cultural diversification of the research landscape.

Regarding the methodologic approach, the analysis confirms the dominance of qualitative and interpretive approaches, including ethnography, visual and arts-based methods, and discourse analysis, used to explore the sensory, spatial, and emotional dimensions of organising. However, more recent studies have begun to employ quantitative and mixed-method designs, including experimental and survey-based approaches (e.g., Ronda & De Gràcia, 2022), indicating an increasing methodological pluralism and a willingness to operationalise aesthetic constructs empirically.

Regarding its evolution, organisational aesthetics is undergoing a digital and hybrid transformation. Post-2020 studies reveal that aesthetic experience increasingly extends beyond physical environments into digital platforms and hybrid workspaces (Tuzcuoğlu et al., 2025; Merriman et al., 2025). Aesthetics now encompasses virtual atmospheres, algorithmic visibility, and camera-mediated performances, reshaping how employees experience embodiment, organizational identity and belonging in hybrid organisations.

As for future research directions, the findings point to five main avenues for development: i) theoretical integration, linking aesthetics to sensemaking, institutional theory, and identity regulation to explain how aesthetic regimes stabilise or challenge organisational norms; ii) ethical and responsible organising aimed at exploring how aesthetic design can support inclusion, diversity, and sustainability (Ivanaj et al., 2018; Ricciardelli et al., 2024); iii) measurement and operationalisation aimed to developing reliable instruments to capture the multidimensional nature of aesthetic organizational experience; iv) digital and hybrid work aesthetics aimed at analysing how technology mediates sensory experience and power relations; v) cross-cultural perspectives aimed at comparing aesthetic sensibilities and practices across cultural contexts to enrich theoretical and practical understanding.

This research shows that organisational aesthetics has matured into a distinct, interdisciplinary, and methodologically diverse field, with expanding theoretical depth and global reach. It highlights that aesthetics is not only about beauty or design but about how organisations feel, look, and make sense of themselves through

sensory and material experiences. For practitioners, this perspective offers valuable insights into how aesthetic awareness can enhance leadership, engagement, and well-being, particularly in digitally mediated work contexts. For scholars, it sets the stage for a new generation of studies that treat aesthetics as a core organisational capability and ethical responsibility, essential for reimagining organisations as human-centred, creative, and regenerative spaces.

7 Ethics Declaration

The research was conducted in accordance with relevant ethical guidelines and principles and official approval was not required.

8 AI Declaration

Artificial Intelligence assistance was used for linguistic proofreading and improvement of English readability. No AI-generated content, ideas, or analysis were included in the scientific development of the paper.

References

Barry, D. and Rerup, C. (2006) 'Going mobile: Aesthetic design considerations for product and organisation', Organization Science, 17(2), pp. 155–165.

Bathurst, R. (2009). Enlivening management practice through aesthetic engagement: Vico, Baumgarten and Kant. *Philosophy of Management*, *7*(2), 61-76.

Bjerke, R., Ind, N. and de Paoli, D. (2010) 'Aesthetics and branding', EuroMed Journal of Business, 5(3), pp. 223–242.

Chang, H.-T., Chou, Y.-J., Miao, M.-C. and Liou, C.-L. (2023) 'On aesthetic preferences', Asia Pacific Management Review, 28(4), pp. 239–250.

Cocks, R., & Reveley, J. (2007). Pragmatism, music and emotion: bridging the Organisational Aesthetics subject-object divide. International Journal of Work Organisation and Emotion, 2(2), 129. https://doi.org/10.1504/ijwoe.2007.017014

Du, W., Ghorbani, M., Ni, Z. and Pan, S.L. (2024) 'Digital platforms and OA', Information Systems Journal, 34(1), pp. 56–78.

Ivanaj, V., Shrivastava, P., & Ivanaj, S. (2018). The value of beauty for organizations. *Journal of Cleaner Production*, *189*, 864-877.

Kachtan, D., & Wasserman, V. (2015). (Un) dressing masculinity: The body as a site of ethno-gendered resistance. *Organization*, *22*(3), 390-417.

Kirillova, K., Fu, X., & Kucukusta, D. (2020). Workplace design and well-being: aesthetic perceptions of hotel employees. *The Service Industries Journal*, *40*(1-2), 27-49.

Lewis, R., & Aune, K. (2023). Aesthetic Labor in Religious Contexts: Women Encountering Modest Dress in the Workplace in the UK and Saudi Arabia. Fashion Theory, 27(5), 709–735. https://doi.org/10.1080/1362704X.2023.2172782

Linstead, S.A. (2018) 'Feeling the reel of the real: Framing the play of film and organisation', Organization Studies, 39(5–6), pp. 667–689.

Marzi, G., Balzano, M., Caputo, A. and Pellegrini, M.M. (2025) 'Guidelines for bibliometric-systematic literature reviews: 10 steps to combine analysis, synthesis and theory development', International Journal of Management Reviews, 27(1), pp. 81–103.

Merriman, K.K., Kostanski, S.A., Ebrahimi, N. and Montag-Smit, T.A. (2025), The Signaling Effect of Biophilic Job Posts: Do Applicants Trade Green for Green?. Hum Resour Manage, 64: 825-839. https://doi.org/10.1002/hrm.22285

Nair S, Hegde N (2026;), "Workplace aesthetics and employee behaviour: introducing the Office Peacocking Scale (OPS)". International Journal of Organizational Analysis, Vol. ahead-of-print No. ahead-of-print. https://doi.org/10.1108/IJOA-07-2025-5716

Raadik, K., Kuura, A. Organisational aesthetics in temporary project organisations: a systematic literature review. Manag Rev Q 75, 2577–2616 (2025). https://doi.org/10.1007/s11301-024-00442-z

Ricciardelli, A., Amoruso, P. and Di Liddo, F. (2024) 'Urban atmospheres and organisational life', Urban Science, 8(1), pp. 12–25.

Ronda, L., & De Gràcia, E. (2022). Does office aesthetics drive job choice? Boosting employee experience and well-being perception through workplace design. *Employee Relations: The International Journal*, *44*(5), 1077-1091.

Sapir, S., Wasserman, V. and Frenkel, M. (2025) 'Constructing sexuality and aesthetics at work', Organization Studies, 46(4), pp. 521–540.

Strati, A. (2019). The riddle and the chair: aesthetics, art and design in organizational life. *Studi di estetica*, *2019*(XLVII, IV serie, 3/2019), 77-100.

Strati, A. (2025) 'Aesthetics and the paradoxical supremacy of ethos in political power', Journal of Political Power, 18(1), pp. 1–18.

Teven, J.J. and Comadena, M.E. (1996) 'Office aesthetic quality and outcomes', Communication Research Reports, 13(2), pp. 190–197.

Tuzcuoğlu, D., Appel-Meulenbroek, R., Borgers, A., Arentze, T., & Sungur, A. (2025). Home or office? The impact of home and office attributes on workplace decisions in hybrid working. *Building and Environment*, 113254.

Vonk, L. (2024). Identity regulation and aesthetic control: The affordances and restraints of uniforms in service work. *Current Sociology*, 00113921241275683.

Wasserman, V. (2012). To be (alike) or not to be (at all): aesthetic isomorphism in organisational spaces. *International Journal of Work Organisation and Emotion*, *4*(1), 22-41.

Wasserman, V., & Frenkel, M. (2015). Spatial Work in Between Glass Ceilings and Glass Walls: Gender-Class Intersectionality and Organizational Aesthetics. Organization Studies, 36(11), 1485-1505.

Warren, S. (2008). Empirical Challenges in Organizational Aesthetics Research: Towards a Sensual Methodology. Organization Studies, 29(4), 559–580. https://doi.org/10.1177/0170840607083104

Witz, A., Warhurst, C. and Nickson, D. (2003) 'The labour of aesthetics and the aesthetics of organisation', Organization, 10(1), pp. 33–54.

AI-Mediated Knowledge Curation in Platform-Based E-Commerce

Stefano Allega, Daniele Allega, Sara El Matouk, Giovanni Spatola
Universitas Mercatorum, Rome, Italy

Abstract

Digital platforms have become entrepreneurial infrastructures rather than simple marketing channels. For start-ups and small and medium-sized enterprises operating in platform-based e-commerce, social media content, online product reviews and customer-service chat transcripts form a distributed textual environment through which opportunities, risks and customer expectations are continuously signalled. This paper develops a conceptual and methodological framework for understanding how artificial intelligence mediates the curation of these multi-source textual assets. The study examines how digital transformation, new technologies and platform dependency reshape entrepreneurial behaviours and knowledge work. Drawing on digital entrepreneurship, the knowledge-based view, platform ecosystem research, the attention economy and emerging literature on AI-mediated knowledge work, the paper argues that AI-enabled analytics operate as a curation layer that extracts, ranks, synthesises and repurposes textual signals. This layer can improve opportunity recognition, experimentation and customer responsiveness, but it also generates a knowledge commoditisation paradox: the same processes that make entrepreneurial knowledge scalable and operationally useful may reduce contextual richness, increase dependency on platform intermediaries and deepen information asymmetries. The paper proposes a mixed-method research design combining semi-structured interviews, digital ethnography, topic modelling, sentiment analysis and intent classification. This design enables researchers to connect entrepreneurs' interpretive practices with the computational patterns found across social, review and support texts. The paper contributes to digital entrepreneurship by conceptualising multi-source text as entrepreneurial knowledge infrastructure; to organisational learning by clarifying how AI-mediated curation converts feedback into routines; to platform economy scholarship by linking visibility, reputation and data access to knowledge governance; and to AI governance by identifying validation routines for trustworthy curation. Managerially, the framework supports SMEs, consultants and platform managers in developing auditable, cross-source and human-accountable knowledge routines. It also suggests policy directions concerning fake reviews, sustainability claims, data portability and platform transparency.

Keywords: Entrepreneurship; Digital Transformation; Entrepreneurial Behaviours; Knowledge Curation; Digital Platforms

Paper type: Academic Research Paper

1 Introduction

Entrepreneurship increasingly occurs inside digital environments in which platforms organise attention, transaction, reputation and learning. E-commerce marketplaces, social media applications and customer-service tools do not merely help firms sell products. They structure what potential customers see, what entrepreneurs can measure, how trust is produced and how quickly a young venture can learn from the market. For start-ups and small and medium-sized enterprises (SMEs), this shift is both enabling and constraining. Platforms reduce entry barriers by providing ready-made infrastructures for discovery, payments, fulfilment and communication. At the same time, they create dependence on algorithmic visibility, platform rules and data-access conditions that entrepreneurs often cannot control.

This paper examines one specific but central dimension of that transformation: AI-mediated knowledge curation in platform-based e-commerce. The empirical phenomenon motivating the paper is simple. Entrepreneurs now face a continuous flow of textual signals distributed across social media comments, influencer posts, online reviews, marketplace questions, chatbot transcripts, complaint messages and product-return explanations. These signals contain knowledge about customer preferences, unmet needs, perceived quality, delivery problems, sustainability expectations, pricing sensitivity and competitor positioning. However, they are noisy, fragmented and unevenly accessible. Artificial intelligence (AI) tools are therefore increasingly used to classify, summarise, rank and transform these texts into actionable insights.

The paper argues that AI-mediated curation should not be understood as a purely technical support activity. It is a knowledge process that shapes entrepreneurial behaviour. When AI systems decide which complaints are urgent, which review themes matter, which social posts indicate emerging demand and which customer-service conversations should be escalated, they influence opportunity recognition, experimentation and resource allocation. They also influence the language through which entrepreneurs describe their markets. Knowledge becomes scalable, but it may also become standardised and detached from the context in which it was produced.

The study therefore asks: how does AI-mediated curation of platform-based textual data reshape entrepreneurial knowledge practices in e-commerce start-ups and SMEs? The paper answers this question conceptually and methodologically. It develops a framework that connects digital entrepreneurship, knowledge-based theory, platform ecosystem research and AI governance. It also proposes a mixed-method research design suitable for future empirical application.

2 Theoretical Background

Digital entrepreneurship refers to entrepreneurial action enabled, constrained or transformed by digital technologies (Nambisan, 2017; von Briel et al., 2018). Unlike traditional venture creation, digital entrepreneurship unfolds in environments where products, processes and market interactions are partly dematerialised. Boundaries between innovation, communication and distribution are blurred. Entrepreneurs can test a product story through a social media post before manufacturing at scale, use marketplace reviews as product-development input and adapt advertising messages through automated analytics. Digital transformation therefore changes not only the tools available to entrepreneurs but also the temporal rhythm and evidentiary basis of entrepreneurial decision-making (Vial, 2019).

Within this environment, knowledge becomes a strategic asset. The knowledge-based view sees firms as institutions for integrating and applying knowledge (Grant, 1996; Kogut and Zander, 1992). Entrepreneurial firms are especially dependent on knowledge because they operate under uncertainty and limited resources. They must interpret weak signals faster than larger competitors while avoiding premature conclusions based on incomplete data. Organisational learning literature further shows that firms transform experience into routines, but those routines can either enable adaptation or lock organisations into narrow interpretations (Cohen and Levinthal, 1990; Nonaka, 1994).

Platform-based e-commerce intensifies this challenge. Platforms act as multi-sided markets that coordinate interactions between sellers, buyers, advertisers, reviewers, logistics providers and service partners (Parker et al., 2016; Gawer, 2021). They also control important conditions of visibility and data access. Entrepreneurial knowledge is therefore not produced in a neutral environment. It is mediated by ranking systems, review architectures, recommendation algorithms, moderation rules and dashboard metrics. A complaint that becomes visible in a review section may shape reputation more strongly than an equivalent complaint expressed privately. A trend detected by a social media algorithm may be amplified because the platform itself rewards certain forms of engagement.

Recent work on AI and knowledge work suggests that machine learning systems increasingly perform tasks that were previously reserved for human analysts, including classification, summarisation, prediction and recommendation (Davenport and Ronanki, 2018; Raisch and Krakowski, 2021). These systems can increase analytical capacity, especially for SMEs that lack large research departments. Yet AI also creates problems of opacity, bias, explainability and accountability (Pasquale, 2015; Floridi and Cowls, 2019). In entrepreneurial contexts, these risks are amplified because decisions are fast, resources are scarce and validation routines may be informal.

The concept of the knowledge commoditisation paradox is useful here. Prior work by Allega and Puzio (2025a) argues that AI-driven value extraction can make knowledge more operationally useful while simultaneously reducing its contextual richness and concentrating power in those who control infrastructures for extraction. This paper transfers that insight from information-intensive organisations to platform-based e-commerce. The same AI routines that help entrepreneurs convert dispersed feedback into decisions may also flatten ambiguity, privilege measurable signals and deepen dependence on platforms and software providers.

3 Conceptual Framework

The proposed framework treats platform-based textual data as entrepreneurial knowledge infrastructure. Infrastructure is not visible only when it fails; it is the background system that enables action. In e-commerce, customer reviews, social media conversations and support transcripts are part of the knowledge infrastructure because they provide recurring evidence about market needs and venture performance. They are not raw truth. They are situated expressions shaped by platform design, customer expectations, cultural norms and strategic behaviour by competitors or influencers.

AI-mediated curation is defined here as the use of computational systems to select, classify, rank, summarise and repurpose textual signals for entrepreneurial decision-making. The definition emphasises mediation because AI does not simply reveal knowledge that already exists. It organises signals according to categories, thresholds and optimisation objectives. For example, an intent-classification model may divide customer messages into delivery, defect, refund and product-use categories. A topic model may group reviews into

themes such as sizing, packaging, durability or sustainability. A generative system may summarise hundreds of comments into a short report. In each case, the system changes the form in which knowledge becomes available.

The framework contains four interacting layers. The first layer is the platform signal layer, which includes social media posts, comments, marketplace reviews, ratings, questions and customer-service conversations. The second layer is the AI curation layer, which performs filtering, classification, sentiment analysis, clustering, summarisation and anomaly detection. The third layer is the entrepreneurial interpretation layer, where founders, managers and employees evaluate AI outputs in relation to strategy, identity, resources and tacit knowledge. The fourth layer is the action and learning layer, where curated insights are converted into product changes, communication choices, pricing experiments, service routines and governance practices.

The key claim is that entrepreneurial value emerges when these layers are connected through validation routines. AI curation without interpretation risks becoming automated noise reduction. Human interpretation without computational support risks being overwhelmed by volume and fragmentation. Learning occurs when entrepreneurs compare AI outputs with lived experience, inspect contradictory evidence, revisit categories and document the decisions that follow. This means that the quality of AI-mediated curation should be assessed not only by technical accuracy but also by its contribution to reflexive entrepreneurial learning.

Three propositions follow. First, AI-mediated curation improves opportunity recognition when it integrates multiple textual sources rather than relying on a single platform metric. Second, AI-mediated curation strengthens entrepreneurial experimentation when outputs are connected to explicit learning routines, such as hypothesis logs or post-experiment reviews. Third, AI-mediated curation increases the risk of knowledge commoditisation when categories, data access and model outputs are controlled by external platforms without transparency or portability.

4 Research Design

The paper proposes a mixed-method design that can be applied to start-ups and SMEs operating in platform-based e-commerce. The design combines qualitative interpretation with computational text analysis. This combination is appropriate because the research object is both social and technical. Entrepreneurial behaviour cannot be inferred from text-mining outputs alone, while interviews alone cannot capture the scale and patterning of digital textual environments.

The qualitative component would include semi-structured interviews with founders, marketing managers, customer-service employees and external consultants. Interviews would explore how firms collect textual data, which AI tools they use, how they validate outputs, how decisions are documented and how platform dependency is perceived. Digital ethnography would complement interviews by observing dashboards, workflows, review-management practices and communication routines. This element is important because entrepreneurs may describe their practices differently from how work is actually organised under time pressure.

The computational component would analyse three families of text: social media content, online reviews and support conversations. Topic modelling can identify recurrent themes across large volumes of text. Sentiment analysis can estimate the emotional direction of customer evaluations, while recognising that sentiment should be treated as an imperfect indicator rather than as an objective measure. Intent classification can distinguish between requests for information, complaints, praise, suggestions and purchase hesitation. Where possible, computational outputs should be compared across platforms to identify whether a theme is platform-specific or genuinely cross-source.

The two components are integrated through iterative comparison. Interview findings inform the categories used in text analysis, and computational results generate questions for follow-up interviews. For example, if topic modelling shows that packaging is a recurrent issue in reviews but entrepreneurs do not mention it as a strategic concern, researchers can investigate whether the signal is ignored, already solved or considered irrelevant. If support transcripts show repeated confusion about product use, but social media content is positive, researchers can examine whether public reputation hides private friction.

Sampling should focus on SMEs and start-ups that sell through at least one digital platform and use at least one AI-enabled tool for analytics, customer support or content management. Cases should vary by sector, firm age, platform dependency and internal analytical capacity. This variation would allow researchers to compare firms that rely almost entirely on external dashboards with firms that build their own cross-source knowledge routines. Ethical safeguards are necessary because customer texts may include personal data, complaints or commercially sensitive information. Anonymisation, access control and clear consent procedures should be incorporated from the beginning of the study.

Table 1: Proposed research design and analytical purpose

Research element	Main data or activity	Analytical purpose
Semi-structured interviews	Founders, managers, customer-service staff and consultants	Understand decision routines, validation practices and perceived platform dependency
Digital ethnography	Dashboards, review workflows and AI-enabled support tools	Observe how curated outputs are embedded in everyday entrepreneurial work
Text analysis	Social media comments, online reviews and support transcripts	Identify topics, sentiment, intent and cross-source patterns
Iterative integration	Comparison between interview themes and computational outputs	Connect human interpretation with machine-readable patterns and contradictions

The table summarises the research design. It is intended to remain editable in the submitted document and to support comparison between human interpretation and computational analysis.

5 Analytical Dimensions and Expected Findings

The framework suggests five analytical dimensions for studying AI-mediated knowledge curation.

The first is source diversity. Entrepreneurial knowledge is stronger when firms compare multiple sources rather than privileging the most visible metric. A high rating on a marketplace may coexist with repeated support complaints. Social media enthusiasm may not translate into repeat purchase. Cross-source comparison protects entrepreneurs from over-learning from a single platform environment.

The second dimension is interpretive ownership. AI systems can generate summaries, but entrepreneurs must decide what those summaries mean for the venture. Interpretive ownership exists when teams understand the categories used by their tools, can challenge outputs and can connect findings to strategic choices. Without this ownership, AI becomes an external authority rather than a learning aid. This risk is especially relevant for SMEs that purchase packaged analytics tools but lack the expertise to inspect their assumptions.

The third dimension is temporal responsiveness. Digital platforms accelerate feedback cycles. Entrepreneurs can detect reactions soon after a campaign, product launch or service failure. AI can increase responsiveness by highlighting anomalies and emerging themes. However, speed can also encourage reactive behaviour. A sudden increase in negative comments may reflect a genuine quality problem, a logistics delay, a viral misunderstanding or coordinated manipulation. Responsiveness must therefore be balanced by verification.

The fourth dimension is platform dependency. Many AI-curated insights are produced from data that entrepreneurs access through platform interfaces. If platforms change their ranking rules, restrict data exports or alter review visibility, firms may lose part of their knowledge base. Dependency is not only commercial; it is epistemic. Entrepreneurs may come to understand their customers through the categories made available by platforms. This can narrow strategic imagination and reinforce short-term optimisation.

The fifth dimension is accountability. AI-mediated curation can influence product design, advertising claims, customer treatment and resource allocation. Firms therefore need routines for documenting how outputs are used. Accountability does not require every entrepreneur to become a data scientist. It requires clear responsibility for data quality, model limitations, human review and decision traceability. Prior research on community feedback and trustworthy AI systems (Allega and Puzio, 2025b; Allega and Puzio, 2025c) is relevant because it shows the importance of preserving disagreement and uncertainty rather than converting all feedback into a single authoritative score.

6 Practical implications

For entrepreneurs and SME managers, the paper suggests that AI-mediated curation should be designed as a learning routine rather than as a dashboard feature. A practical routine may begin with mapping all available textual sources, including public reviews, private support messages, social comments and return explanations. Firms can then define a limited set of decision categories, such as product improvement, service failure, communication clarity, pricing concern and opportunity signal. The aim is not to automate all interpretation, but to make interpretation more systematic.

Second, SMEs should use cross-source validation. A theme should be treated as strategically important when it appears across several sources or when it is strongly connected to a critical business objective. For example, a sustainability claim that generates positive social engagement but sceptical reviews may require evidence-based clarification before being amplified. A packaging complaint that appears in support transcripts and reviews may deserve operational attention even if average ratings remain high.

Third, entrepreneurs should maintain human accountability for AI outputs. Teams can create short decision logs indicating what the AI system suggested, what evidence was inspected, what action was taken and what result followed. This practice is simple but powerful. It transforms AI from a black-box recommendation engine into part of an organisational memory. It also supports future auditing if decisions are challenged by customers, partners or regulators.

Fourth, consultants, accountants and digital advisors serving SMEs can play an important role. Many small firms do not have internal data governance expertise. Advisors can help entrepreneurs choose tools, define categories, document processes and avoid over-reliance on platform dashboards. This is consistent with the broader role of knowledge professionals in supporting digital transformation among SMEs.

Finally, platform managers and policymakers should recognise that data access and transparency affect entrepreneurial knowledge. Review authenticity, data portability, explainable ranking changes and access to historical performance data are not only technical issues. They shape the ability of SMEs to learn, compete and innovate. Policies that support fair access to platform-generated knowledge can therefore contribute to a healthier entrepreneurial ecosystem.

7 Discussion

The central insight of the paper is that AI-mediated knowledge curation is an entrepreneurial capability. It combines technological tools, interpretive judgement and governance routines. This perspective moves beyond two common simplifications. The first simplification is technological optimism, which treats AI analytics as a straightforward solution to information overload. The second is technological pessimism, which treats AI as necessarily alienating or manipulative. The more useful position is relational: AI becomes valuable or harmful depending on how it is embedded in entrepreneurial practice.

The framework also clarifies how the knowledge commoditisation paradox operates in platform-based e-commerce. AI systems make knowledge more mobile, comparable and operational. A founder can receive a summary of thousands of comments in minutes. A marketing manager can compare sentiment across channels. A customer-service team can identify recurring problems. These are real benefits. Yet the same processes may detach signals from context. Sarcasm, cultural references, platform-specific norms and minority concerns can disappear in aggregated summaries. What is easy to count may dominate what is strategically important.

For theory, the paper contributes to digital entrepreneurship by showing that entrepreneurial opportunity is increasingly mediated by textual infrastructures and AI categorisation. It contributes to the knowledge-based view by identifying curation as a mechanism through which dispersed external knowledge is transformed into internal routines. It contributes to platform ecosystem research by emphasising epistemic dependency: platforms do not only mediate transactions but also mediate what entrepreneurs can know. It contributes to AI governance by bringing accountability and validation into the everyday practices of SMEs, rather than treating governance as a concern only for large technology firms.

The paper has limitations. It is conceptual and methodological rather than empirical. Its propositions require testing across sectors and countries. E-commerce SMEs differ in their analytical maturity, customer base, regulation and dependence on platforms. Future research should examine longitudinal cases to understand how curation routines evolve as firms grow. Comparative studies could also examine whether platform-native ventures behave differently from traditional SMEs undergoing digital transformation.

8 Conclusions

This paper developed a framework for understanding AI-mediated knowledge curation in platform-based e-commerce. It argued that social media content, online reviews and customer-service conversations form a distributed knowledge infrastructure for start-ups and SMEs. AI tools can help entrepreneurs extract and organise signals from this infrastructure, supporting opportunity recognition, experimentation and responsiveness. However, these tools also create risks of standardisation, opacity and platform dependency.

The proposed framework connects platform signals, AI curation, entrepreneurial interpretation and organisational learning. Its main practical message is that entrepreneurs should not delegate knowledge work to AI systems. They should build accountable routines that combine computational scale with human judgement. The main theoretical message is that digital entrepreneurship increasingly depends on the governance of knowledge flows, not only on access to technologies. New technologies matter most when they change what entrepreneurs notice, how they learn and how they act under uncertainty.

8 Acknowledgements

The authors acknowledge the intellectual continuity between this paper and prior research on AI-driven knowledge commoditisation, trustworthy artificial intelligence architectures and community feedback mechanisms. No external funding was received for the preparation of this paper.

9 Ethics Declaration

This paper is a conceptual and methodological academic research paper based on secondary literature and a proposed future research design. No primary data from human participants, customers or firms were collected for the preparation of this manuscript. Ethical clearance was therefore not required at this stage. Any future empirical implementation involving interviews, customer-service transcripts, social media data or online reviews should obtain appropriate institutional approval, informed consent where applicable, anonymise personal data and comply with relevant data-protection rules.

10 AI Declaration

AI tools were used to support language drafting, structural editing and formatting of this manuscript. The authors reviewed, selected, revised and verified the resulting material, remain fully responsible for the scholarly content, and confirm that AI-generated text was not used as a substitute for authorial judgement, source evaluation or ethical accountability.

References

Allega, D. (2025) The artificial legal intelligence society as an open, multi-sided platform for law-as-computation. In Panait, M., Rădulescu, I., Popescu, C. and Voica, M.C. (Eds.), Creativity and Innovation in Digital Economy: Book of Abstracts, Petroleum-Gas University of Ploiești Publishing House, pp 136-138.

Allega, D., Puzio, G. and Rizzo, G. (2026) The artificial legal intelligence society as an open, multi-sided platform for law-as-computation. In Panait, M., Rădulescu, I.G., Tudorică, B., Popescu, C., Voica, M.C. and Nica, M. (Eds.), Proceedings of the 4th International Conference Creativity and Innovation in Digital Economy, Petroleum-Gas University of Ploiești Publishing House, pp 66-75.

Allega, D., Puzio, G. and Rizzo, G. (2026) The knowledge commoditization paradox: Theoretical and practical challenges of AI-driven value extraction in information-intensive organizations. In Panait, M., Rădulescu, I.G., Tudorică, B., Popescu, C., Voica, M.C. and Nica, M. (Eds.), Proceedings of the 4th International Conference Creativity and Innovation in Digital Economy, Petroleum-Gas University of Ploiești Publishing House, pp 95-107.

Allega, D., Puzio, G. and Rizzo, G. (2026) MERL-T: A multi-expert architecture for trustworthy artificial legal intelligence. In Panait, M., Rădulescu, I.G., Tudorică, B., Popescu, C., Voica, M.C. and Nica, M. (Eds.), Proceedings of the 4th International Conference Creativity and Innovation in Digital Economy, Petroleum-Gas University of Ploiești Publishing House, pp 76-83.

Allega, D., Puzio, G. and Rizzo, G. (2026) Reinforcement learning from community feedback: A novel framework for artificial intelligence in social science domains. In Panait, M., Rădulescu, I.G., Tudorică, B., Popescu, C., Voica, M.C. and Nica, M. (Eds.), Proceedings of the 4th International Conference Creativity and Innovation in Digital Economy, Petroleum-Gas University of Ploiești Publishing House, pp 84-93.

Allega, D. and Puzio, G. (2025a) The knowledge commoditization paradox: Theoretical and practical challenges of AI-driven value extraction in information-intensive organizations. In Panait, M., Rădulescu, I., Popescu, C. and Voica, M.C. (Eds.), Creativity and Innovation in Digital Economy 2025: Book of Abstracts, Petroleum-Gas University of Ploiești Publishing House, pp 66-68.

Allega, D. and Puzio, G. (2025b) MERL-T: A multi-expert architecture for trustworthy artificial legal intelligence. In Panait, M., Rădulescu, I., Popescu, C. and Voica, M.C. (Eds.), Creativity and Innovation in Digital Economy 2025: Book of Abstracts, Petroleum-Gas University of Ploiești Publishing House, pp 170-171.

Allega, D. and Puzio, G. (2025c) Reinforcement learning from community feedback: A novel framework for artificial intelligence in social science domains. In Panait, M., Rădulescu, I., Popescu, C. and Voica, M.C. (Eds.), Creativity and Innovation in Digital Economy 2025: Book of Abstracts, Petroleum-Gas University of Ploiești Publishing House, pp 92-94.

Cohen, W.M. and Levinthal, D.A. (1990) Absorptive capacity: A new perspective on learning and innovation, Administrative Science Quarterly, Vol. 35, No. 1, pp 128-152.

Davenport, T.H. and Ronanki, R. (2018) Artificial intelligence for the real world, Harvard Business Review, Vol. 96, No. 1, pp 108-116.

Floridi, L. and Cowls, J. (2019) A unified framework of five principles for AI in society, Harvard Data Science Review, Vol. 1, No. 1.

Gawer, A. (2021) Digital platforms and ecosystems: Remarks on the dominant organizational forms of the digital age, Innovation, Vol. 24, No. 1, pp 110-124.

Grant, R.M. (1996) Toward a knowledge-based theory of the firm, Strategic Management Journal, Vol. 17, Winter Special Issue, pp 109-122.

Kogut, B. and Zander, U. (1992) Knowledge of the firm, combinative capabilities, and the replication of technology, Organization Science, Vol. 3, No. 3, pp 383-397.

Nambisan, S. (2017) Digital entrepreneurship: Toward a digital technology perspective of entrepreneurship, Entrepreneurship Theory and Practice, Vol. 41, No. 6, pp 1029-1055.

Nonaka, I. (1994) A dynamic theory of organizational knowledge creation, Organization Science, Vol. 5, No. 1, pp 14-37.

Parker, G.G., Van Alstyne, M.W. and Choudary, S.P. (2016) Platform Revolution: How Networked Markets Are Transforming the Economy and How to Make Them Work for You, W.W. Norton, New York.

Pasquale, F. (2015) The Black Box Society: The Secret Algorithms That Control Money and Information, Harvard University Press, Cambridge, MA.

Raisch, S. and Krakowski, S. (2021) Artificial intelligence and management: The automation-augmentation paradox, Academy of Management Review, Vol. 46, No. 1, pp 192-210.

Vial, G. (2019) Understanding digital transformation: A review and a research agenda, Journal of Strategic Information Systems, Vol. 28, No. 2, pp 118-144.

von Briel, F., Davidsson, P. and Recker, J. (2018) Digital technologies as external enablers of new venture creation in the IT hardware sector, Entrepreneurship Theory and Practice, Vol. 42, No. 1, pp 47-69.

Zahra, S.A., Liu, W. and Si, S. (2023) How digital technology promotes entrepreneurship in ecosystems, Technovation, Vol. 119, Article 102457.

Smart Healthy Cities Advancing

Mauro Romanelli
University of Naples Parthenope, Napoli, Italy

Abstract

Cities are adopting a smart city strategy in order to support and improve smart urban health to achieve better quality of life within smart healthy urban communities. Cities are rediscovering the smart urban health as a driver of urban growth and innovation. Cities are going into the future, encouraging urban stakeholders' engagement and collaboration, and embracing a smart city mind set and urban health view to urban value creation. The study aims to elucidate how the European cities are rethinking on the future of urban communities, adopting a smart view and focusing on urban health to improve the wealth of urban communities. The study relies on qualitative research analysis. The research analyzes two case studies drawn from policy documents related to European cities that are rethinking on future design of the city, and planning a smart view to urban growth, leading to smart health as a driver of urban innovation. Brussels and Wien have designed a smart healthy city strategy, promoting urban health, and driving smart urban innovation. Smart urban health planning helps shape sustainable and wealthy cities, leading to responsible urban innovation and sustainable growth, supporting multi-actor collaborative innovation, and focusing on community development. Smart urban health and innovation support collaborative spaces in order to drive smart and responsible urban innovation. Introducing the smart urban health view within urban planning is enriching the approach to smart city vision, enabling the smart mind set as a way to design and implement organizational and collaborative spaces for innovation. Driving smart urban health refers to collaborative frameworks that involve a variety of urban stakeholders and actors that contribute to shaping the city a smart innovative community which is able to contribute to sustainable growth and social innovation. The smart city vision supports a collaborative platform that enhances smart health, leading to social innovation within urban communities.

Keywords: smart city, smart health, healthy city, urban innovation.

Paper type: Academic Research Paper

1 Introduction

Today, cities are designing smart city plans to drive responsible urban innovation, leading to *more than human* and post-anthropocentric healthy cities (Kocak and Sancaktutan, 2025), generating desired outcomes for humans and nonhumans (Yigitcanlar, 2018; Yigitcanlar *et al.*, 2021), meeting urban healthcare as a grand societal challenge (Voegtlin, Scherer, Stahl and Hawn, 2022; Agarwal *et al.*, 2025), and empowering healthcare innovation, leading to high wellbeing of citizens (Alves *et al.*, 2025), and supporting collaborative frameworks for quality of life improvement (Mora *et al.*, 2023; Yigitcanlar *et al.*, 2021).

Smart cities contribute to supporting community wellbeing (Thompson *et al.*, 2023), empowering the e-health in order to improve public health, bridging digital technologies and business (Eysenbach, 2001).

A healthy city improves its environments and expands its resources, supporting people empowerment and engagement, achieving their highest potential (WHO, 2022). With more than 50 percent of people living in cities, providing services-solutions to make the city wealthy and innovative, healthy and sustainable relies on smart city development to improve urban health innovation, wellbeing and quality of life, leading to collaborative approaches (Thompson *et al.*, 2023; Meijer and Bolivar, 2016).

As organizational spaces (Knox, 2010), net action (Czarniawska, 2002), cities are rethinking on smart city urban planning towards truly smart urban healthy system of systems, doing good, leading to responsible urban innovation, towards an organised maturity through open health services and platforms that enable positive freedoms (Lee *et al.*, 2014; Yigitcanlar, 2018; Scherer, Neesham, Schoeneborn and Scholz, 2023).

Urban health matters critically, whereas smart mind set helps improve urban health innovation and ubiquitous services for citizen wellbeing, quality of life, safety, and nature friendliness, following a collaborative model (Sapienza *et al.*, 2022; Ji *et al.*, 2021; Meijer and Bolivar, 2016).

Smart and healthy cities rely on using technology, community and policy as key drivers for responsible urban health management, design and deployment, city operational design and innovation, leading to equitable and living cities, focusing on wellbeing and healthcare issues and benefits (Trencher and Karvonen, 2019; Yigitcanlar *et al.*, 2021; Jahroh *et al.*, 2025).

Still few studies focus on smart urban health planning and responsible innovation (Solanas *et al.*, 2014). How do cities aim at planning smart and healthy cities? How smart city view helps support smart healthy city and responsible urban innovation? This study aims to elucidate how cities and municipalities of Brussels and Wien

are dealing with urban health management and innovation, adopting a smart city plan. After the introduction, the theoretical background relates to smart and healthy cities that contribute to responsible urban innovation. Methodological section follows in the third section. Findings are presented in the fourth section. Finally, discussion and conclusions are outlined.

2 Smart and Healthy Cities for Responsible Urban Innovation

Smart healthy cities contribute to digital urban ecosystems for wealth and wellbeing, empowering open and collaborative ways for health services innovation to improve healthy human and urban life (Paskaleva, Evans and Watson, 2021), leading to sustainable healthy development through technology- community-policy drivers (Yigitcanlar and Cugurullo, 2020).

Smart city is a healthy system of systems which is able to generate desired outcomes and futures for all humans-and-non-humans (Yigitcanlar, 2018). Smart city development leads to quality of life, supporting community integration, relying on environmental and healthy wellbeing (Macke *et al.*, 2018), fostering collaboration among technology developers, healthcare providers and policymakers to create an inclusive healthcare system (Agarwal *et al.*, 2025). Smart city improves healthcare as a critical infrastructure (Washburn *et al.*, 2010).

Smart health aims to promote high position of health within society, augmenting mobile health with sensing capabilities of smart city, leading to a convergent ubiquitous health paradigm (Solanas *et al.*, 2014). Healthy smart cities provide an urban healthy place for citizen engagement and cooperation (Kocak and Sancaktutan, 2025) in order to reduce the impact of pollution, carbon emission, waste and heat, improving human health. Healthy city is a system of health sectors (Ziafati, Bafarasati and Sharifi, 2024), improving health equity and well-being, and meeting urban health as grand societal challenge and driver (Vardoulakis and Kinney, 2019).

Smart city vision and implementation should develop as responsible urban innovation, leading to trustworthy healthcare innovation and technology for users' care acceptance, aligning technology with societal and human values within urban communities (Li *et al.*, 2023).

Cities aim at planning smart city progressive cities, dealing with social challenges, and shaping urban collaborative ecosystems for responsible innovation (Mora *et al.*, 2023), leading to a smart community that enables a *continuum* of public-private actors which combine technology and policy, and shape the urban community, improving social and healthcare services, and leading to responsible urban health innovation (Gil-Garcia, Pardo and Nam, 2015; Meijer and Bolivar, 2016; Voegtlin, Scherer, Stahl and Hawn, 2022).

A smart healthy city aims at combining technology, policy and community-driven sources to provide innovative as well as collaborative initiatives that contribute to engendering and improving health care services that benefit urban ecosystems.

Rethinking on the urban planning following a smart city view helps shape smart responsible, healthy and sustainable cities (Yigitcanlar, 2018), following a socio-technical and collaborative view (Meijer and Thaens, 2018), meeting the health needs of urban communities, tackling social challenges, focusing on collaborative frameworks and community engagement for urban health innovation and wellbeing (Trencher and Karvonen, 2019).

Smart healthy cities rely on collaborative and community-driven platforms that support urban stakeholder communication, providing digital solutions for quality of life, health and wellbeing improvement, embedding health smart city model for smart healthy and wealthy communities (Mora, Deakin and Reid, 2019; Mouton *et al.*, 2019).

3 Methodological Section

The study has employed a qualitative and descriptive methodology. Brussels and Wien have clearly identified a smart city planning to make a smart healthy city. Rethinking smart healthy responsible urban innovation leads to *smart and healthy more than human* cities, and relies on information sources collection drawn by city planning to track the pathway towards a smart innovative responsible city (Yigitcanlar, 2018; Cristiano and Zilio, 2021; Thompson *et al.*, 2023; Sancino and Hudson, 2020). The study focuses on two illustrative and comparative case studies to explore a comprehensive understanding of a phenomenon still in progress, emerging in contemporary times, whereas the boundaries between the phenomenon and the context are not clearly evident following a case study methodology through an empirical inquiry without the rigidity of a predefined structure for observations and analysis (Yin, 2009). The research methods involved case selection and data collection from policy documents (*Smart Climate City Strategy Vienna; Smart City Strategy of the City of Brussels*).

4 Findings

The *Smart City Strategy of the City of Brussels* frames urban development around the integration of digital innovation, sustainability, and citizen well-being, positioning smart urban health as a central pillar of its broader vision. The strategy adopts a holistic understanding of health that extends beyond healthcare services to encompass environmental quality, social inclusion, mobility, and digital accessibility. In this sense, Brussels aligns with the principles of the *Healthy City* approach, which emphasizes the social determinants of health, equitable access to resources, and the creation of environments that support physical, mental, and social well-being.

Within the strategy, smart urban health is conceptualized as the use of data, digital tools, and intelligent infrastructures to monitor, understand, and improve the health of the urban population. This includes the deployment of sensors to track air quality, noise levels, and urban heat; the use of open data platforms to support evidence-based policymaking; and the promotion of digital health services that enhance accessibility and efficiency. The city also prioritizes active mobility, green spaces, and sustainable transport systems, recognizing their direct impact on public health and quality of life. The Healthy City dimension is reflected in Brussels' commitment to equity, participation, and social cohesion. The strategy highlights the need to reduce health disparities by ensuring that technological innovation benefits all residents, including vulnerable groups. Citizen participation is encouraged through co-creation platforms, digital engagement tools, and community-based initiatives that allow residents to contribute to decision-making processes. This participatory governance model reinforces the idea that health is a shared responsibility shaped by collective action. By integrating the Healthy City framework with smart technologies, Brussels advances toward a *Smart Healthy City* model. This hybrid approach leverages digital innovation not as an end in itself but as a means to enhance well-being, sustainability, and inclusiveness. The strategy positions Brussels as a city where technology supports healthier lifestyles, more resilient environments, and a governance system that is transparent, collaborative, and oriented toward long-term public value.

A healthy city adopts the ambitions of the 2030 Energy-Climate Plan and reduces its total energy consumption and its greenhouse gas emissions, in particular by focusing on the production of renewable energy, and puts the health of its inhabitants first, by improving the quality of the air in the city and by combatting noise pollution in public spaces and residential areas. The Smart City strategy contributes to improving waste management and the cleanliness of public spaces through digital technologies and data analysis; managing exemplary buildings by exploiting new technologies for the recovery of primary energy resources, by providing monitoring tools to objectify and improve the reduction of energy and water consumption in the territory, and by promoting the energy management of municipal buildings, integrating sustainable development into the city's processes by developing the capacity to continuously monitor the territory, using technologies to make urban spaces greener and friendly for biodiversity.

The *Smart Climate City Strategy of Vienna* positions health, well-being, and climate resilience as central pillars of the city's long-term urban transformation, integrating the principles of *smart urban health* and the *Healthy City* approach into a comprehensive framework that links climate action, digital innovation, and social equity. Vienna's vision recognizes that urban health is shaped by both healthcare systems and environmental quality, mobility patterns, housing conditions, and social cohesion. *Smart urban health* relies on using data, digital tools, and intelligent infrastructures to monitor and improve environmental and health-related indicators, deploying sensor networks to track air quality, noise pollution, heat stress, and mobility flows, enabling evidence-based interventions that enhance public health. Digital platforms and open data systems support transparency and allow policymakers, researchers, and citizens to access real-time information. Vienna also promotes digital health services, telemedicine, and smart care solutions to improve accessibility and efficiency, particularly for vulnerable populations. The ambition and aims of a smart city plan is to safeguard health's interests and wellbeing of people, helping shape safe and clean urban environments that minimize waste and reduce water and soil pollution, noise and heat pollution as far as possible. Smart city framework relies on promoting health and social inclusion by supporting policies that contribute to improving the health of urban community and residents, so achieving and meeting social and health goals, and helping Vienna's healthcare sector to contribute significantly to decarbonisation, focusing strongly on sustainable urban development issues, improving urban health as social innovation, and digital and social inclusion, whereas urban health system improvement is leading to local economic development too. Vienna advances toward a *Smart Healthy City* model merging climate action, digital innovation, and health-oriented urban planning, reducing emissions, expanding green infrastructure, and improving energy efficiency directly enhance public well-being. Technology is as a means to create a resilient, inclusive, and health-promoting urban environment. *Smart Climate City Strategy* demonstrates how cities can integrate smart technologies with social and environmental policies to support healthier, more sustainable, and more equitable urban futures.

5 Discussion and Conclusions

The European cities are planning smart urban health for quality of life improvement, leading to smart responsible and healthy urban communities, engines of innovation and collaborative open spaces. Smart city vision develops the city of the future and supports responsible innovation to achieve sustainable urban, leading to healthy and safe urban environments. Smart urban health planning helps to shape sustainable and wealthy cities. Advancing smart urban health helps drive responsible urban innovation and sustainable urban growth, leading to multi-actor collaborative innovation, focusing on community development, leading to smart city as a platform for responsible collaborative innovation. The Wien smart vision relies on the role of technology to support inclusive healthy and wealthy smart urban ecosystems by enhancing co-participation, and strengthening co-decision-making for urban responsible innovation and sustainable growth, meeting grand urban challenges, dealing with climate change and pollution issues. The city and the municipality of Brussels aims at embedding both health and smart vision in the urban planning, looking at the goal of sustainable and more than human city, fostering a collaborative and inclusive approach in order to shape urban meeting places for urban health and value creation. Cities are planning a smart city design to improve urban health and wellbeing of citizens, providing a community-led digital platform that helps shape wealthy urban communities, making a healthy responsible urban community, an engine of innovative services and approaches through healthcare, citizens and urban stakeholders. Smart city development relies on promoting responsible innovation by developing socially inclusive and open health urban ecosystems that enhance the interaction that concerns humans, technologies and community, leading to healthier cities on becoming intelligent cities. Cities are adopting a smart city vision in order to improve urban health competences and issues. The future and maturity of smart city development relies on embedding health care services and innovation within the smart city ecosystem, leading to a smart and healthy more-than-human community, and meeting urban health, a grand societal opportunity, and combining technology, smart city policy and community to make smart urban health a responsible innovation, and develop both a mind set and way leading to wealthy urban ecosystems.

6 Ethics Declaration

I declare that any ethical clearance was required for the research.

7 AI Declaration

I declare that AI tools were not used to create this paper.

References

Agarwal, V., Benmamoun, Z., Anjum, M., Mathiyazhagan, K., Akim, M. and Pironti, M. (2025) "Urban Innovation Dilemmas: Tackling the Challenges for Urban Growth in Smart City", *Technology in Society*, 103194.

Alves, V., Marreiros, G., Fdez-Riverola, F., Neves, J., Ribeiro, J. and Vicente, H. (2024) Advancing Smart Urban Futures in the Healthcare Sector, *International Symposium on Distributed Computing and Artificial Intelligence*, Springer Nature Switzerland, Cham, pp 210-221.

Cristiano, S. and Zilio, S. (2021) "Whose health in whose city? A systems thinking approach to support and evaluate plans, policies, and strategies for lasting urban health", *Sustainability*, Vol. *13, No.* 21, 12225.

Czarniawska, B. (2002) "Remembering while Forgetting: The Role of Automorphism in City Management in Warsaw", *Public Administration*, Vol. 62, No. 2, pp 163-173.

Eysenbach, G. (2001) "What is e-health?", *Journal of Medical Internet Research*, 3.2

Gil-Garcia, J.R., Pardo, T. A. and Nam, T. (2015) "What makes a city smart? Identifying core components and proposing an integrative and comprehensive conceptualization", *Information Polity*, Vol. 20, No. 1, pp 61-87.

Jahroh, S., Indrawan, D., Junaid, Z. B., Abdullah, A., Fahmi, I. and Siddique, M. (2025) "The smart and healthy city business model Canvas—A post Covid-19 resilience for smart city business modeling framework", *Clinical Ehealth*, Vol. 8, pp 78-93.

Ji, T., Chen, J. H., Wei, H.H. and Su, Y. C. (2021) "Towards people-centric smart city development: Investigating the citizens' preferences and perceptions about smart-city services in Taiwan", *Sustainable Cities and Society*, Vol. 67, 102691.

Kocak, B. and Sancaktutan, E. (2025) "Healthy Cities and World Health Organization Healthy Cities Project", *Cities*, Vol. 163, 106043.

Knox, H. (2010) "Cities and organisation: The information city and urban form", *Culture and Organization*, Vol. 16, No. 3, pp 185-195.

Lee, J. H., Hancock, M. G. and Hu, M. C. (2014) "Towards an effective framework for building smart cities: Lessons from Seoul and San Francisco", *Technological Forecasting and Social Change*, Vol. 89, pp 80-99.

Li, W., Yigitcanlar, T., Browne, W. and Nili, A. (2023) "The making of responsible innovation and technology: An overview and framework", *Smart Cities*, Vol. 6, No. 4, pp 1996-2034.

Macke, J., Casagrande, R.M., Sarate, J.A.R. and Silva, K.A. (2018) "Smart city and quality of life: Citizens' perception in a Brazilian case study", *Journal of Cleaner Production*, Vol. 182, pp 717-726.

Meijer, A. and Bolívar, M.P.R. (2016) "Governing the smart city: a review of the literature on smart urban governance", *International review of administrative sciences*, Vol. 82, No. 2, pp 392-408.

Meijer, A.J. and Thaens, M. (2018) "Urban Technological Innovation: Developing and Testing a Sociotechnical Framework for Studying Smart City Projects", *Urban Affairs Review*, Vol. 54, No. 2, pp 363-387.

Mora, L., Appio, F.P., Foss, N. J., Arellano-Gault, D. and Zhang, X. (2023) "Organizing for Smart City Development: Research at the crossroads. Introduction to the Special Issue", *Organization Studies*, Vol. 44, No. 10, pp 1559-1575.

Mora, L., Deakin, M. and Reid, A. (2019) "Strategic principles for smart city development: A multiple case study analysis of European best practices", *Technological Forecasting and Social Change*, Vol. 142, pp 70-97.

Mouton, M., Ducey, A., Green, J., Hardcastle, L., Hoffman, S., Leslie, M. and Rock, M. (2019) "Towards 'smart cities' as 'healthy cities': health equity in a digital age", *Canadian Journal of Public Health*, Vol. *110*, No. 3, pp 331-334.

Paskaleva, K., Evans, J. and Watson, K. (2021) "Co-producing smart cities: A Quadruple Helix approach to assessment", *European Urban and Regional Studies*, Vol. 28, No. 4, pp 395-412.

Sancino, A. and Hudson, L. (2020) "Leadership in, of, and for smart cities–case studies from Europe, America, and Australia", *Public Management Review*, Vol. 22, No. 5, pp 701-725.

Sapienza, M., Nurchis, M. C., Riccardi, M. T., Bouland, C., Jevtić, M. and Damiani, G. (2022) "The adoption of digital technologies and artificial intelligence in urban health: a scoping review", *Sustainability*, Vol. 14, No. 12, 7480.

Scherer, A.G., Neesham, C., Schoeneborn, D. and Scholz, M. (2023) "New challenges to the enlightenment: How twenty-first-century sociotechnological systems facilitate organized immaturity and how to counteract it", *Business Ethics Quarterly*, Vol. 33, No. 3, pp 409-439.

Thompson, S., Rahmat, H., Marshall, N., Steinmetz-Weiss, C., Bishop, K., Corkery, L., ... and Tietz, C. (2023) "Merging smart and healthy cities to support community wellbeing and social connection", *Encyclopedia*, Vol. *3*, No. 3, pp 1067-1084.

Trencher, G. and Karvonen, A. (2019) "Stretching "smart": Advancing health and well-being through the smart city agenda", *Local Environment*, Vol 24, No. 7, pp 610-627.

Vardoulakis, S. and Kinney, P. (2019) "Grand challenges in sustainable cities and health", *Frontiers in Sustainable Cities*, Vol. 1, 7.

Voegtlin, C., Scherer, A. G., Stahl, G. K. and Hawn, O. (2022) "Grand societal challenges and responsible innovation", *Journal of Management Studies*, Vol. *59*, No. 1, pp 1-28.

Washburn D, Sindhu U, Balaouras S, Dines RA, Hayes NM and Nelson LE (2010) Helping CIOs Understand 'Smart City' Initiatives: Defining the Smart City, its Drivers, and the Role of the CIO, Forrester Research, Cambridge, MA

WHO (2022) How to develop and sustain healthy cities in 20 steps. WHO European Healthy Cities Network. WHO European Healthy Cities Network.

Yigitcanlar, T. and Cugurullo, F. (2020) "The sustainability of artificial intelligence: An urbanistic viewpoint from the lens of smart and sustainable cities", *Sustainability*, Vol. 12, No. 20, 8548.

Yigitcanlar, T. (2018) "Smart city policies revisited: Considerations for a truly smart and sustainable urbanism practice", *World Technopolis Review*, Vol. 7, No. 2, pp 97-112.

Yigitcanlar, T., Corchado, J. M., Mehmood, R., Li, R. Y. M., Mossberger, K. and Desouza, K. (2021) "Responsible urban innovation with local government artificial intelligence (AI): A conceptual framework and research agenda", *Journal of Open Innovation: Technology, Market, and Complexity*, Vol. 7, No. 1, 71.

Yin, R.K. (2009) How to do better case studies, *The SAGE handbook of applied social research methods*, *2*, pp 254-282.

Ziafati Bafarasat, A. and Sharifi, A. (2024) "How to achieve a healthy city: a scoping review with ten city examples", *Journal of Urban Health*, Vol. 101, No. 1, pp 120-140.

Cultural Organizations and Young Audiences: A Collaborative View through Artificial Intelligence

Mauro Romanelli[1], Alexandra Zbuchea[2], Monica Bira[2]
[1] University of Naples Parthenope, Napoli, Italy
[2] National School of Political and Administrative Studies, Romania, Bucharest

Abstract

Cultural organizations are developing and implementing Artificial Intelligence (AI) and digital interactive technologies in order to improve cultural and social value production, meeting and engaging the young audiences in cultural heritage landscape. AI is increasingly shaping the strategies and operations of cultural organizations such as museums, galleries, libraries, and arts institutions. AI helps cultural organizations to improve the interaction with young people who act as agents of cultural heritage. AI helps enable cultural and heritage institutions and organizations as collaborative spaces that function as socio technical environments where communities, professionals, and technologies interact to co-create cultural value. AI supports cultural organizations to shape cultural heritage environments as collaborative spaces and meeting places for social innovation development. AI aims to facilitate collaborative interpretation and contribute to cultural knowledge. The study aims at investigating the relationships between cultural organizations, artificial intelligence and young audiences. AI supports cultural heritage sites as collaborative spaces for social innovation. AI helps drive cultural organizations to play a significant role in cultural value and knowledge creation, acting as agents of social change and value. Engaging young audiences in becoming proactive actors who contribute to social value creation within cultural heritage ecosystems supports social and technological innovation within cultural organizations that are rediscovering new sources and energies in order to behave as social value-driven and knowledge-led intelligent organizations that contribute to public value and better quality of life, driving social change within communities. The challenge of cultural organizations is to use AI to engage young people and make them proactive actors and responsible co-producers of social and cultural value. Cultural organizations enable youth to shape narratives and contribute to increasing the youth engagement, positioning themselves as platforms of social exchange and dialogue, enabling cultural heritage sites as digital and onsite meeting collaborative places that contribute to supporting cultural value creation.

Keywords: artificial intelligence, cultural organizations, collaborative spaces.

Paper type: Academic Research Paper

1 Introduction

Digital interactive technologies and Artificial Intelligence (AI) are leading museums and cultural organizations into the future. Cultural organizations are providing new services and facilities that contribute to meeting young audiences, enriching the experiences of visitors, and restructuring processes within the cultural organizations' landscape, by developing, using and implementing AI within cultural heritage.

AI is providing significant opportunities for cultural organizations which aim to innovate and expand their reach. AI-driven tools are expanding how cultural organizations interact with audiences. The growing integration of AI within processes and strategies of cultural organizations and museums is reshaping how institutions curate, interpret, and communicate heritage, with particular implications for younger audiences, who increasingly expect digital interactivity, personalization, and participatory experiences. Cultural organizations are using AI to increasingly engage young audience. In particular, AI technologies exert influence on behaviors and lifestyles, trust and perceived authenticity of young persons (Lata and Rana, 2025; Bachnik et al., 2025; Chen et al., 2026). Engaging youth in the cultural heritage management processes helps preserve the transmission of heritage values, and supports the diversification of knowledge systems, leading to inclusive approaches in heritage management (Zhang, 2025). AI generative systems facilitate collaborative annotation, storytelling, and interpretation, enabling users with diverse backgrounds to contribute meaningfully to cultural knowledge.

AI refers to a broad set of digital technologies. AI facilitates the automation of existing processes hitherto carried out by other means (human labour, and other forms of calculation) (Caramiaux, 2023). On the other side, youth, aged between 15-24 are the decision-makers of the future and agents of cultural heritage (Zhang, 2025).

AI is transforming cultural organizations into responsive, data-driven, socially innovative and participatory ecosystems, particularly attractive to younger audiences who seek personalization, interactivity, and hybrid experiences. AI helps cultural organizations and audiences to experiment cultural heritage landscape as a meeting as well as collaborative and participatory space for innovation (Montanari, Mattarelli, and Scapolan, 2020; Capdevila, 2019; Zhang, 2025).

Few studies investigate the relationships between cultural organizations that introducing and implementing AI technologies and applications, and the young visitors and audiences. Few studies investigate how the interaction between AI and cultural organizations is opening to new opportunities of engendering collaborative spaces for innovation within the cultural heritage landscape.

The study is exploratory and qualitative. The aim of the study is to elucidate how cultural organizations are managing AI technologies and engaging new audiences, involving young visitors, rediscovering new sources, energies and capabilities, leading to collaborative spaces and innovation within cultural heritage environments.

The adoption of AI is driving broader organizational and social change and innovation. Cultural institutions are increasingly required to develop new skills and infrastructures, leading to the importance of interdisciplinary collaboration between technologists, curators, and policymakers, and overcoming the resistance to change and the lack of technical expertise as barriers that especially influence social performances of cultural organizations.

The study is organized as follows. After the introduction the theoretical background on the relationships between cultural organizations and AI with regards to collaborative spaces is presented. In the third paragraph the methodological section follows. The fourth section presents the relationship between AI, cultural organizations and the young audiences. Finally, discussion and conclusions are outlined.

2 Towards Collaborative Spaces through Cultural Organizations Managing Artificial Intelligence

Collaborative and open spaces provide meeting places for social interaction, enabling several actors to construct a pathway for social innovation and value creation (Montanari, Mattarelli, and Scapolan, 2020). Collaborative spaces support multi-actor and co-initiation of collaborative innovation processes (Sørensen and Torfing, 2018, and provide opportunities for the development of creative solutions through the collective knowledge and experience of actors that contribute to innovation genesis (Hartley, Sørensen and Torfing, 2013), leading to collaborative working practices and spaces (Yacoub and Hefliger, 2024), that enhance creative platforms (Capdevila, 2019), and foster openness, collaboration and knowledge sharing and skills (Boutillier *et al.*, 2020).

Participatory and collaborative spaces in cultural organizations are increasingly recognized as key infrastructures for enabling participatory innovation, collective knowledge production, and digitally augmented cultural engagement of youth, and supporting participatory approaches (Zhang, 2025). AI may play a significant role in augmenting the potential of cultural heritage environments as collaborative as well as meeting places for social exchange and new knowledge construction. In particular, the practices of cultural heritage practitioners, archivists, and librarians rely on collaboration and cooperative practices. They jointly contribute to generating an intelligence which is artificial by sense making what they interpret in their collections (Caramiaux, 2023).

Technology enables collaboration between users and cultural context (Giugliano and Laudante, 2020). Recent scholarship emphasizes the relationship between co-creation and participatory initiatives as improved by the use of AI with regards to young visitors' experiences. Engaging and involving the young visitors through AI within cultural heritage value creation is opening new ways to make cultural heritage ecosystems as collaborative as well as participatory meeting places for innovation. AI helps explore collaborative design methods to achieve human-computer interaction and improve cultural landscape design (Tang, Zheng, Chen, Yan and Chen, 2024). Emerging AI digital technologies contribute to promoting the development of participation, and creating new opportunities for young people who engage in various digitally mediated activities, offering new spaces for empowering the voices of young people, and providing new spaces of dialogue, sense making, knowledge and information exchange, and social value. AI supports current heritage management practices, driving the young people to act as proactive co-producers and intelligent and aware agents of meaning construction, abandoning the actual position of merely users and passive recipients of heritage values (Zhang, 2025). AI helps enable cultural and heritage institutions and organizations to shape collaborative spaces that function as socio technical environments where communities, professionals, and technologies interact to co create cultural value (Simon, 2010). AI can facilitate such participation by analyzing user-generated content, clustering themes, or generating adaptive displays that reflect collective contributions. Moreover, AI supports institutional decision-making by analyzing visitor flows, sentiment, and engagement patterns, enabling museums to design youth-centered strategies grounded in empirical evidence (Bonacini and Giaccone, 2022).

3 Methodological Section

The study employed a qualitative, descriptive and exploratory methodology to analyse how the interplay between the AI technologies, cultural organization, and the audience, in particular, the young audiences and visitors enable the rise of collaborative as well as organisational spaces that contribute to improving the quality of visitors' experiences, and engendering new ways and practices of providing cultural heritage services, and

driving value creation through participatory as well as cooperative ways, leading to culture-led and creative-driven itineraries for value creation. The study is conceptual, exploratory and theoretical because the boundaries between cultural organizations, AI and the collaborative ways to organisational innovation within cultural heritage ecosystems are not yet defined in a satisfying way. Conceptual papers aim to propose new frameworks with regards to untested relationships among constructs, focusing on integration and association of logical arguments, without providing data (Gilson and Goldberg, 2015). The study refers to the analysis of articles related to the relationship between cultural organizations, AI and young audiences. The selected contributions are interpreted in a narrative synthesis to elucidate new perspectives and advance theoretical frameworks on emerging issues (Denyer and Tranfield, 2006; Dixon-Woods *et al.*, 2004).

4 Artificial Intelligence, Cultural Organizations and Young Audiences

The academic literature is increasingly concerned with how cultural organizations use digital technologies to improve their offerings and attract more customers. In parallel, the same technologies, including AI, are shaping new ways of cultural engagement, especially in the case of young persons, the digital-native ones. In the context of a relatively low interest in cultural consumption, the use of digital and AI technologies could make cultural organizations more appealing, including by offering new experiences.

Technology is credited as a factor of resilience in cultural sector (Wang and Huang, 2024). It contributes to diversification of the offer and even with systemic changes in the way culture is produced, distributed and consumed (Yordanova and Todorova, 2025). Adoption of digital and AI technologies is not simple, especially for cultural organizations due to various types of limits – budgetary, lack of expertise, willingness for change etc. Additionally, these changes are not always seen as positive considering various stakeholders.

Technological innovation takes many forms. Yordanova and Todorova (2025) identified nine main ones: Artificial Intelligence (AI), Robotics, Big Data, Virtual Reality (VR), Augmented Reality (AR), Digitalization, 3D Printing, Data Analytics, and the Internet of Things (IoT). Their analysis identified the most used ones as AI, digitalization and VR. These might be more common considering their direct positive impact both on cultural organizations and cultural consumption.

Especially the AI integration offers not only a technological improvement, but also a shift in how culture is created, mediated and experienced. Young audience, already more literate in terms of digital and AI technologies, are the primary beneficiaries and users of AI integrated into cultural organizations and cultural consumption. Even several years ago, AI – e.g. the chat bot – was credited with mediating heritage and engaging audiences (Gaia, Boiano and Borda, 2019). The developments in the field of generative AI have led to new ways of personalized and meaningful experience of heritage and culture (Trichopoulos, Ordoumpozanis and Caridakis, 2026).

AI is credited with significantly influencing both the consumption behaviour and the lifestyle of young persons, especially (Lata and Rana, 2025). For instance, the interest and support towards more sustainable behaviour among young people is related to increased use of AI. Young people tend to have a complex relationship with AI in terms of trust and perceived authenticity (Bachnik et al., 2025; Chen et al., 2026). Studies show that young people make a critical but imperfect effort to discern what is “real” and “authentic” when AI is involved. This means that cultural organizations have to be aware that young audiences are influenced by AI and open to it, but in the same time they can approach it with some caution and display selective trust. This means cultural organizations need to be transparent about AI use and intentional in framing authenticity related to it.

The perceived authenticity also influences emotional engagement (Tang, Dong and Cheng, 2025). Studies show that authenticity in AI-driven experiences is critical for memory retention and identity formation among young persons. Still, cultural characteristics and values influence the way humans related to AI-experiences and contribute to contextualizing the social and cultural impact (Shin, Al-Imamy and Hwang, 2022).

Cultural organizations could use AI in many ways to engage meaningfully the audience, especially the young one. Chat bots and conversational agents could contribute to personalized assistance and gamification (Iodice, Clemente and Bifulco, 2024). AI, in general, contributes to compelling storytelling, contributing to cultural inclusiveness, engagement, enhancing cultural inclusivenes, sensitivity and understanding (European Parliament, 2023; Lau, 2025; UNESCO, 2022; Wang and Zhang, 2026). AI can also be used to curate more appealing and relevant exhibitions, considering specific young audiences (Trichopoulos, Ordoumpozanis and Caridakis, 2026). AR and VR technologies enable immersive and multisensory experienced, enhance emotional perception as well as learning (Tsoukala et al., 2024; Xu et al., 2025).

AI has numerous implications in other aspects of cultural and heritage management, such as preservation of heritage, featuring both positive and negative aspects (Ghaith and Hutson, 2024). Considering its many advantages, for many cultural organizations, AI becomes a must in order to attract younger audiences (Vidu,

Zbuchea and Pinzaru, 2021), but also to contribute to more effective heritage and cultural management. Critical to its sustainable impact is, still, the establishment of ethical practices in this field.

5 Discussion and Conclusions

There is an increasing interest within cultural organization which are using AI and intelligent technologies to improve cultural services and engage the young audiences, and focus actions and policies to involve the young generations. Intelligent technologies are reshaping the relationships between cultural organizations and its social and cultural ecosystems, providing opportunities for cultural value creation, audience engagement, and organizational innovation. Cultural organizations take the opportunity of developing and implementing AI technologies in order to assume the responsibility to promote inclusivity and support the participation and the engagement of the audience in cultural value production. In particular, AI helps cultural organizations to engage more actively young visitors, enabling them to gather information and reflect on cultural heritage meaning and importance in order to support organizations and people to advance dealing with complexity, and promoting social issues.

Increasingly, advanced, digital and interactive information and communication technologies exert influence on social and organizational change within cultural organizations that aim at dealing with technology to drive organizational as well as strategic innovation. Moreover, unexplored territories for knowledge augmentation and learning are emerging whereas the intelligent technologies support processes of knowledge creation and sharing, and contribute to fostering the interaction and communication between audiences and the professionals at work within cultural organizations. AI technologies are entering into the cultural heritage field, opening new opportunities for building cultural collaborative spaces and engendering new authentic cultural value. Cultural organizations are becoming meeting places that empower cooperative efforts and energies, involving a variety of actors (human, collective, technological) that contribute to improving the visitors' experiences, and producing social and cultural value, leading to social innovation and inclusion issues.

There are social, managerial and organizational implications emerging from this preliminary investigation. Cultural organizations contribute to enhancing social value, and shaping proactive young visitors who rediscover the potential of bridging human capabilities and technological infrastructure. Introducing and implementing AI within cultural organizations' landscape enables management processes, redesigning organizational and human resources management policies. Making cultural heritage sites as collaborative meeting places for interaction between the audience and cultural organizations is opening new perspectives, and searching for management and organizational solutions.

The sustainable adoption of AI relies on balancing innovation with ethical responsibility, ensuring that AI enhances cultural understanding without compromising authenticity, inclusivity, or human agency. As cultural organizations continue to evolve, AI will likely play an increasingly central role in shaping how young people encounter, interpret, and co-create cultural heritage.

Cultural organizations are dealing with AI design and implementation in order to engage new audiences and restore the relationship with young visitors, leading to intelligent organizations which rely on knowledge-led sources and capabilities.

Future research perspective imply to explore long term AI impact, with regards to enhancing cultural diversity, and understanding the potential and the evolving role of human expertise in AI-supported environments. Research agenda relies on testing the perceptions of young visitors in using AI applications to manage the cultural heritage sources, and appreciate how cultural organizations are engaging them, meeting their needs and preferences.

Further research also aims to identify the trajectories of change and innovation provided by the impact of AI technologies within the organization of cultural institutions, leading to a variety of potential organizational solutions that contribute to value creation, and bridging social, organizational and technological infrastructures.

6 Ethics Declaration

I declare that any ethical clearance was required for the research.

7 AI Declaration

I declare that AI tools were not used to create this paper.

References

Bachnik, K., Nowacki, R., Fandrejewska-Nowakowska, A., Zatwarnicka-Madura, B. and Eisenbardt, T. (2025) "AI-Generated Advertising Through the Eyes of Generation Z", *Contemporary Economics*, Vol. 19, No. 4, 417.

Bonacini, E. and Giaccone, S.C. (2022) "Gamification and cultural institutions in cultural heritage promotion: a successful example from Italy", *Cultural trends*, Vol. 31, No. 1, pp 3-22.

Boutillier, S. Capdevila, I. Dupont, L. and Morel, L. (2020) "Collaborative spaces promoting creativity and innovation", *Journal of Innovation Economics & Management*, Vol. 31, No. 1, pp 1-9.

Capdevila, I. (2019) "Motivating participation in collaborative spaces", *Journal of Business Strategy*, Vol. 40, No. 2, pp 14-21.

Caramiaux, B. (2023) "AI with museums and cultural heritage", *AI in Museums*, pp 117-130.

Chen, N., Lai, Z., Liu, Y., Li, J., Wang, R. and Yan, P. (2026) "See, trust, and interact: how AI disclosure shapes high school students' trust", *Information Research an International Electronic Journal*, *31*(iConf), pp 1099-1145.

Denyer, D. and Tranfield, D. (2006) Using Qualitative Research Synthesis to Build an Actionable Knowledge Base, *Management Decision*, Vol. 24, pp 213-227.

Dixon-Woods, M., Agarwall, S., Young, B., Jones, D. and Sutton, A. (2004) *Integrative Approaches to Qualitative and Quantitative Evidence*, Health Development Agency, London, www.hda.nhs.uk

European Parliament (2023) *Artificial intelligence in the context of cultural heritage and museums: Complex challenges and new opportunities* (EPRS Briefing No. PE 747.120). European Parliamentary Research Service. https://www.europarl.europa.eu/RegData/etudes/BRIE/2023/747120/EPRS_BRI(2023)747120_EN.pdf

Gilson, L.L. and Goldberg, C.B. (2015) "Editors' comment: so, what is a conceptual paper?", *Group & Organization Management*, Vol. 40, No. 2, pp 127-130.

Gaia, G., Boiano, S. and Borda, A. (2019) "Engaging museum visitors with AI: The case of chatbots". In: *Museums and Digital Culture: New Perspectives and Research*. Springer International Publishing, pp 309-329.

Ghaith, K. and Hutson, J. (2024) "A qualitative study on the integration of artificial intelligence in cultural heritage conservation", *Metaverse*, Vol. 5, No. 2, 2654.

Giugliano, G. and Laudante, E. (2020) Design as collaborative connection between user, technology and cultural context. In *IOP Conference Series: Materials Science and Engineering*, IOP Publishing, Vol. 949, No. 1, p 012010

Hartley, J., Sørensen, E. and Torfing, J. (2013) "Collaborative innovation: A viable alternative to market competition and organizational entrepreneurship", *Public administration review*, Vol. 73, No. 6, pp 821-830.

Iodice, G., Clemente, L. and Bifulco, F. (2024) "The role of chatbot technology in the cultural value proposition: A managerial perspective in human-like interaction". In: *The use of artificial intelligence in digital marketing: Competitive strategies and tactics*, IGI Global Scientific Publishing, pp 166-187.

Lata, S. and Rana, K. (2025) "AI's influence on young consumer behavior: fostering sustainable consumption", *Young Consumers*, Vol. 26, No. 5, pp 848-864.

Lau, K.H.C. (2025). Designing and Evaluating Gen-AI for Cultural Resilience. In *Proceedings of the 27th International Conference on Multimodal Interaction*, pp 706-710.

Montanari, F. Mattarelli, E. and Scapolan, A.C. (2020) *Collaborative spaces at work: Innovation, creativity and relations*, Routledge, London.

Shin, D., Al-Imamy, S. and Hwang, Y. (2022) "Cross-cultural differences in information processing of chatbot journalism: chatbot news service as a cultural artefact", *Cross Cultural & Strategic Management*, Vol. 29, No. 3, pp 618-638.

Simon, N. (2010) The participatory museum. Museum 2.0.

Sørensen, E. and Torfing, J. (2018) "Co-initiation of collaborative innovation in urban spaces", *Urban Affairs Review*, Vol. 54, No. 2, pp 388-418.

Tang, M., Dong, J. and Cheng, S. (2025) "Artificial Intelligence Enhanced Integration of Art Education and Cultural Tourism for Student Engagement Improvement through PLS-SEM and IPMA". In: *Proceedings of the 2025 International Conference on AI-enabled Education*, pp 391-396.

Tang, Q., Zheng, L., Chen, Y., Yan, L. and Chen, J. (2024) "Artificial intelligence empowering museum space layout design: Insights from China", *Plos one*, Vol. 19, No. 11, e0310594.

Trichopoulos, G., Ordoumpozanis, K. and Caridakis, G. (2026) "An evaluation of LLM-based chatbots for enhancing the visitor's user experience at cultural exhibits", *ACM Journal on Computing and Cultural Heritage*, Vol. 19, No. 1, pp 1-25.

Tsoukala, S., Kalliampakou, I., Theodoropoulou, A. and Giannoukou, I. (2024) "Digital engagement in cultural institutions: Enhancing visitor experience through innovative technologies". In: *International conference of the international association of cultural and digital tourism*. Springer Nature Switzerland, pp 375-412.

UNESCO. (2022) *Recommendation on the ethics of artificial intelligence*. United Nations Educational, Scientific and Cultural Organization.

Vidu, C., Zbuchea, A. and Pînzaru, F. (2021) Old meets new: integrating Artificial Intelligence in museums' management practices. In: *Strategica. Shaping the Future of Business and Econ*omy, Tritonic, Bucarest, pp 830-844.

Wang, M. and Huang, Y. (2024) "A digital Technology–Cultural resource strategy to drive innovation in cultural industries: A dynamic analysis based on machine learning", *Technology in Society*, Vol. 77, 102590.

Wang, K. and Zhang, S. (2026) "AI-guided and human-guided tours in museum educational tourism: mechanisms of teenagers' long-term memory formation", *Tourism Review*, pp 1-24.

Xu, Z., Dahlan, N. D., Syed-Abdullah, S. I. S., Sahrir, S. and Cao, K. (2025). Augmented Reality for Cultural Preservation: A Case Study of Shaanxi Weinan Shadow Puppetry. In: *Proceedings of the 2025 2nd International Conference on Informatics Education and Computer Technology Applications*, pp 253-260.

Yacoub, G. and Haefliger, S. (2024) “Coworking spaces and collaborative practices”, *Organization*, Vol. 31, No. 1, pp 87-114.
Yordanova, Z. and Todorova, Z. (2025) “Technological innovation in cultural organizations: A review and conceptual mapping framework”, *Digital*, Vol. 5, No. 4, 54.
Zhang, Y. (2025) Engaging Youth with Cultural Heritage in the Digital Age: An Integrated Framework for Youth Digital Participation in Cultural Heritage Management.

The Digital Platform as Cognitive Infrastructure: Investigating Human-AI Synergy and Algorithmic Governance in Investment Management

Paola Demartini, Alessandro Salvatori
Roma Tre University, Rome, Italy

Abstract

Over the past two decades, the financial sector has undergone a profound transformation, firmly positioning information as its primary competitive asset and placing Artificial Intelligence (AI) at the core of this shift. By interpreting diverse data volumes at speeds beyond human capacity, AI serves as a critical supportive technology. Within this landscape, digital platforms like BlackRock's Aladdin have emerged as essential cognitive infrastructures functioning as comprehensive operating systems for modern finance. These platforms seamlessly integrate the entire investment lifecycle into a unified data layer, reshaping organizational roles and business models by facilitating a vital interaction between human insight and algorithmic intelligence.

Despite AI's well-documented technical capabilities, a critical gap remains in understanding the socio-technical dynamics of this human-AI synergy within institutional settings. Current literature predominantly focuses on the predictive accuracy of "black box" models or macro-level regulatory frameworks, overlooking how professionals navigate daily trade-offs between algorithmic performance and explainability. Furthermore, there is limited empirical exploration of the modern financial analyst's "hybrid profile," which requires combining computational literacy with ethical sensitivity within established platform constraints.

To address this gap, this study investigates how digital investment platforms facilitate human-AI synergy while managing structural risks of algorithmic opacity and bias. Employing a qualitative, exploratory approach, the research utilizes semi-structured interviews with senior European Union professionals involved in investment products, digital financial advisory, platform delivery, and risk-related decision-making. The investigation focuses on integrating AI-driven insights into fundamental analysis, perceptions of algorithmic opacity, the EU AI Act's impact on governance, and platforms' role in standardizing information.

Preliminary findings indicate that rather than replacing human experience, AI fundamentally reconfigures it, shifting professional judgment toward problem definition, exception management, and robustness testing. Integrated platforms provide a "common data language" that aligns internal functions, significantly reducing manual errors and operational risk. However, a persistent tension exists between maximizing predictive power and maintaining the transparency demanded by the EU AI Act, which frames AI as a corporate governance issue rather than mere efficiency. Ultimately, the study highlights a "new responsibility pact" requiring auditable decision-making, while warning that "algorithmic convergence" through dominant platforms could dangerously amplify market volatility. Sustainable innovation requires these cognitive infrastructures to be supported by robust data governance.

Keywords: Digital Platform, investment decision, financial sector, BlackRock's Aladdin

Paper type: Academic Research

1 Introduction

Over the past two decades, the global financial sector has undergone a profound structural and epistemological transformation, firmly positioning information processing as its primary competitive asset (Varshney & Varshney, 2026). Within this highly complex and historically intermediated environment, Artificial Intelligence (AI) has emerged not merely as a supplementary tool for operational efficiency, but as a foundational "cognitive infrastructure" (Davenport et al., 2020). AI architectures are now capable of collecting, cleaning, connecting, and interpreting increasingly diverse volumes of both structured and unstructured data, offering insights and executing decisions at speeds that vastly exceed human cognitive capacity (Chen, Härdle, & Ong, 2021; Cao, 2024). This evolution marks a transition from simple digitalization—where physical infrastructure and paper-based systems were translated into digital formats—to an era of intelligent finance, where advanced algorithms dynamically estimate conditional probabilities and integrate textual, visual, and behavioral data into unified analytical pipelines (Arner, Barberis, & Buckley, 2016; Brynjolfsson & McAfee, 2017).

In this landscape, digital investment platforms have emerged as essential mechanisms that facilitate the critical interaction between human insight and algorithmic intelligence (FT, 2020). A preeminent manifestation of this shift is BlackRock's Aladdin platform, which functions as a comprehensive "operating system" for modern finance (BlackRock, 2024). Aladdin and similar platforms seamlessly integrate the entire investment lifecycle—

encompassing portfolio construction, trade execution, risk control, accounting, and regulatory reporting—into a unified, standardized data layer (Buehlmaier & Whited, 2023). As organizations strive for long-term sustainability and resilience in globalized capital markets, these technological platforms represent a new "operational language" that fundamentally reshapes professional roles, required skill sets, and institutional business models (FT, 2020).

Despite the extensive and well-documented technical capabilities of Artificial Intelligence in finance—such as its capacity for high-dimensional pattern recognition, non-linear forecasting, and anomaly detection (Ruff et al., 2021)—a critical gap persists in the academic literature regarding the socio-technical dynamics of this transformation. Current scholarship predominantly bifurcates into two distinct domains: technical evaluations focused on the predictive accuracy of specific "black box" machine learning models (Goodfellow, Bengio, & Courville, 2016), and macro-level legal analyses concentrated on regulatory frameworks, such as the European Union's Artificial Intelligence Act (European Commission, 2023). Consequently, there is a pronounced lack of empirical investigation into the daily realities of "human-AI synergy" within institutional investment settings. The literature often overlooks how financial professionals navigate the continuous, complex trade-offs between maximizing algorithmic predictive performance and satisfying the institutional mandates for explainability and transparency (Barredo Arrieta et al., 2020; Quinn, 2023). Furthermore, there remains limited exploration of the modern financial analyst's "hybrid profile," which necessitates a delicate balance of computational literacy, economic intuition, and ethical sensitivity, all functioning within the rigid constraints of established platform architectures.

To address these theoretical and empirical gaps, this research investigates the socio-technical integration of advanced algorithms in investment banking. The primary research question guiding this study is: *How do digital investment platforms facilitate human-AI synergy while managing the structural risks of algorithmic opacity and bias?* By analyzing AI not merely as an isolated technological deployment but as a systemic architectural shift, this paper seeks to elucidate how financial judgment is being transformed from a purely human cognitive exercise into a technologically mediated, probabilistic discipline (Lopez de Prado, 2020).

2 Research Design and Methodology

To examine Artificial Intelligence as a structural reconfiguration of financial intermediation rather than a mere technical enhancement, this study adopts a qualitative, exploratory, and interpretative methodological framework. The epistemological positioning of this research is intentionally systemic; rather than testing a narrow econometric hypothesis regarding algorithmic generation, the objective is to investigate how AI operates as a cognitive infrastructure that reshapes decision-making processes, governance structures, and competitive dynamics.

The research employs a multi-layered methodological strategy that integrates conceptual-theoretical literature, institutional-regulatory analyses, and empirical-operational investigations. The foundational layer consists of a critical literature review encompassing peer-reviewed academic articles, regulatory frameworks (e.g., the EU AI Act, EBA guidelines), and institutional reports (BIS, OECD, FSB) published predominantly between 2016 and 2025. This documentary analysis serves to construct a theoretical baseline regarding the promises and limitations of financial AI (FSB, 2024; OECD, 2023).

The core empirical investigation utilizes a qualitative case study design centered on BlackRock's Aladdin platform, supplemented by comparative evaluations of platforms such as VantagePoint AI, Tickeron, Danelfin and the Numerai ecosystem. Aladdin was selected as a paradigmatic case due to its systemic relevance, overseeing tens of trillions of dollars in global assets, its hybrid architecture combining deterministic quantitative models with generative AI interfaces, and its widespread cross-institutional adoption (BlackRock, 2024; Financial Times, 2020). To bridge the theoretical framework with operational reality, the study incorporates semi-structured interviews with senior financial professionals operating within the European Union. Participants include senior professionals involved in investment products and services, digital financial advisory, platform delivery, and risk-related decision-making processes (see table 1).

Table 1**:** Research data source Synopsis

Source Category	Number	Selection Criteria	Role in Analysis
Peer-reviewed academic articles	24	Relevance to AI, ML, finance (2016–2025)	Theoretical framework
Institutional reports (BIS, FSB, OECD, EU, ECB)	13	Regulatory authority and policy relevance	Governance & systemic risk analysis
Industry reports (McKinsey, Deloitte, PwC, BlackRock)	17	Practical implementation insight	Operational & strategic assessment

Source Category	Number	Selection Criteria	Role in Analysis
Legal and regulatory documents (AI Act, GDPR, EBA Guidelines)	6	Normative binding frameworks	Compliance dimension
Web-based supplementary sources (Sitography)	36	Definitional clarification or descriptive background	Supporting descriptive context
Case study platforms analysed	6 (Aladdin VantagePoint AI, Tickeron, Danelfin and the Numerai ecosystem)	Systemic significance	Infrastructure evaluation
Semi-structured interviews	3	Direct institutional AI involvement	Empirical validation
Total documentary sources	100	Multi-layer screening	Integrated synthesis

A critical empirical focal point is the implementation of the Aladdin platform within "Omega Bank," a major European banking group. The semi-structured interviews with Omega Bank's head of investment products and the head of the Digital Financial Advisor office were designed to uncover the strategic rationale for adoption, the practical configuration of the platform, and its daily utilization in managing hundreds of thousands of financial instruments (see the questionnaire in annex 1). The interview transcripts were analyzed using a thematic coding methodology to identify recurring operational patterns, discrepancies between academic theory and institutional practice, and the practical challenges of algorithmic governance. While the geographical concentration on the European context and the reliance on a specific institutional case study pose inherent limitations to generalizability, this qualitative approach provides unparalleled depth in interpreting the socio-technical realities of AI integration.

3 Findings

The empirical investigation yields multi-dimensional findings regarding how AI operates beyond current standard implementations, fundamentally restructuring the epistemological basis of investment banking. The findings are organized according to the core thematic branches identified in the conceptual mind map structuring the advanced applications of AI in finance.

3.1 Cognitive infrastructures and the "Common Data Language"

A foundational finding is that the true transformational power of platforms like Aladdin does not stem from autonomous "artificial intelligence" making unilateral decisions, but rather from the establishment of a "common data language" (BlackRock, 2024). Within Omega Bank, Aladdin normalizes disparate data streams related to positions, transactions, market volatility, and macroeconomics into a unified, centralized data layer. This architectural unification ensures that front-office portfolio managers, middle-office risk officers, and back-office compliance teams operate from an identical informational base (Figure 1).

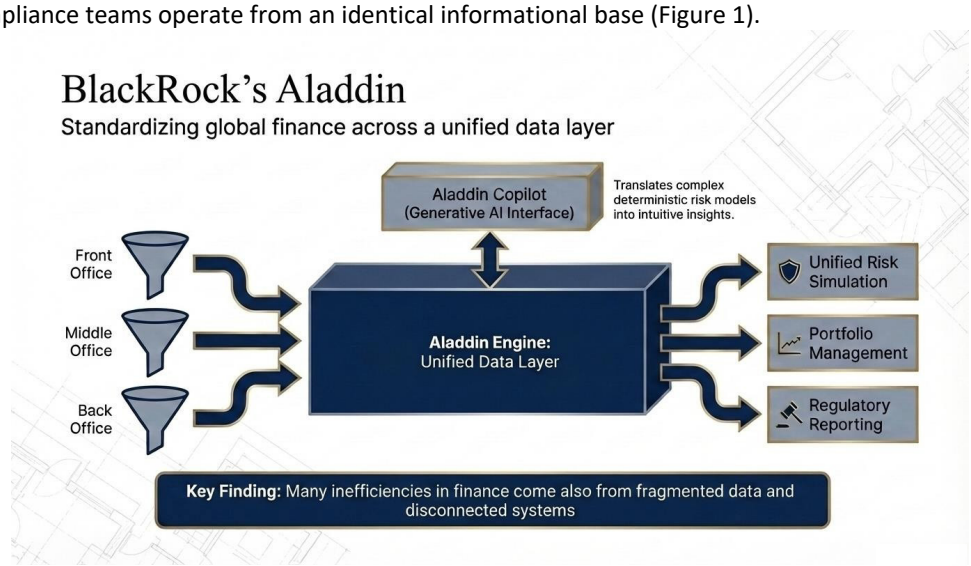


Figure 1: Aladdin engine

By centralizing data on approximately 70,000 financial instruments from client portfolios, within a broader universe of hundreds of thousands of mapped instruments, the platform drastically reduces operational risks historically caused by manual reconciliations, data silos, and informational asymmetries. Consequently, complex analytical tasks—such as simulating the impact of a 200-basis-point interest rate hike or conducting multi-factor stress tests—become almost instantaneous. The platform acts as an enabling infrastructure for organizational standardization, aligning individual discretionary advisory proposals with institutionally mandated model portfolios (Figure 2).

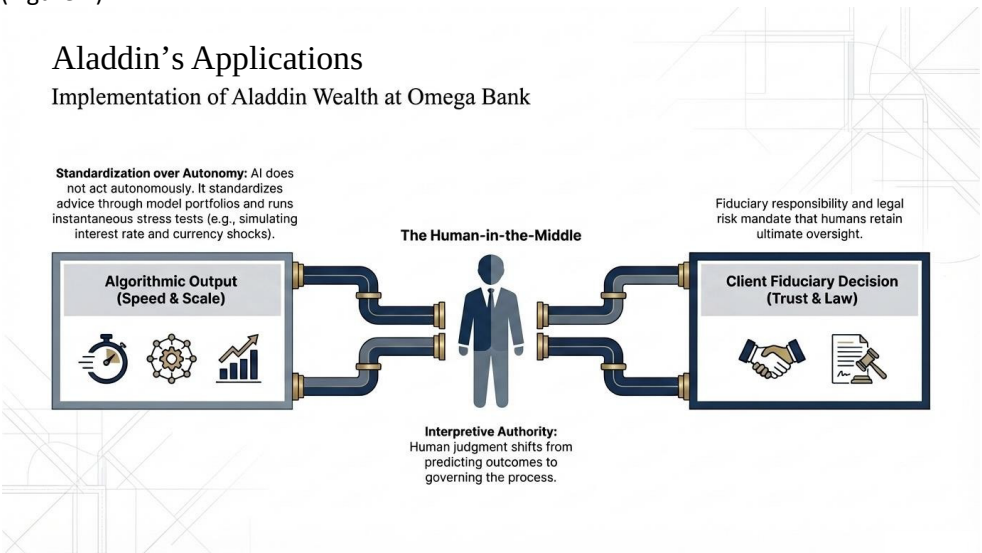


Figure 2: Implementation of Aladdin wealth at Omega bank

3.2 Predicting asset returns: A methodological shift

Moving beyond operational efficiency, the analysis of AI's advanced capabilities reveals a profound methodological shift in the attempt to predict asset returns. For decades, traditional econometric forecasting relied on linear models that imposed restrictive assumptions regarding market stationarity, variable independence, and distributional normality. The findings indicate that AI initiates a paradigm shift from linear modelling to non-linear machine learning, embracing high-dimensional function approximation (Feng, He, & Polson, 2018). AI systems do not approach prediction as the deterministic forecasting of a single future price point; instead, they operate on the principle of *probabilistic inference under uncertainty* (Lopez de Prado, 2020). Modern algorithms process massive datasets to approximate complex conditional probability distributions—estimating the likelihood of specific market regimes, the probability of excess returns, or the trajectory of expected volatility. This shift aligns with the realization that uncertainty in financial markets cannot be eliminated, but it can be mathematically modelled and probabilistically managed.

3.3 Key AI architectures in predictive analytics

The research identifies four key AI architectures driving this methodological shift. First, Deep Learning and Long Short-Term Memory (LSTM) networks are increasingly utilized to capture sequential, long-term dependencies in time-series data (Heaton, Polson, & Witte, 2016). LSTMs are particularly effective at identifying momentum effects and delayed macroeconomic transmission mechanisms across varying market regimes. Second, Classical Machine Learning models, such as Support Vector Machines (SVMs), Random Forests, and gradient-boosting algorithms (e.g., XGBoost), remain vital for cross-sectional stock selection due to their robustness in handling noisy datasets and their relative interpretability compared to deep neural networks (Nakamura, 2022). Third, Natural Language Processing (NLP) and sentiment analysis represent a breakthrough in converting qualitative information—such as central bank press releases, corporate earnings calls, and social media dynamics—into quantifiable metrics that capture behavioral market distortions (Tetlock, 2007; Malo et al., 2014). Finally, the deployment of Hybrid Systems allows institutions to merge traditional financial data with alternative data streams, such as satellite imagery of retail parking lots or mobile geolocation data, expanding the informational surface available for algorithmic analysis (McKinsey & Company, 2023).

3.4 Real-World forecasting platforms

The democratization of AI technologies has led to the proliferation of Real-World Forecasting Platforms, which operationalize these complex architectures into accessible, commercial interfaces. The study examined

several dominant paradigms. *VantagePoint AI* utilizes neural networks grounded in intermarket analysis, processing dozens of correlated variables across commodities, currencies, and equities to generate short-term directional forecasts (VantagePoint, 2025). *Tickeron* aggregates pattern recognition algorithms and probabilistic trend classifiers into a unified terminal, offering users dynamic probabilities of trend continuation or breakout events (Tickeron, 2025). *Danelfin* introduces an innovative cross-sectional prediction model, synthesizing thousands of fundamental and technical indicators through ensemble learning to assign an "AI Score" (from 1 to 10) representing a stock's probability of outperforming a benchmark (Danelfin, 2025). While these platforms compress immense informational complexity into actionable insights, they also risk fostering behavioral overconfidence if users misinterpret high probabilistic scores as deterministic guarantees, particularly given inherent cognitive biases (Kahneman & Tversky, 1979).

3.5 AI-Native organizations: The Numerai paradigm

Beyond retail platforms, the findings highlight the emergence of AI-Native Organizations, which completely restructure the investment production function around algorithmic prediction. *Numerai* represents a radical paradigm shift: it operates as a crowdsourced hedge fund utilizing tournament logic (Numerai, 2026). Numerai provides global data scientists with highly engineered, *obfuscated data*—intentionally stripping the economic semantics from the variables to prevent narrative-driven overfitting. Participants train diverse machine learning models and submit their predictions, which the platform aggregates. Crucially, Numerai introduces an economic incentive design via *staking and burn mechanisms* utilizing its proprietary cryptocurrency (figure 3). If models fail out-of-sample, the participant's staked capital is burned; if successful, they are rewarded (Numerai Docs, 2025). This architecture resolves the issue of accountability in crowdsourcing, acting as an endogenous filter for trust and robustness. Similarly, *WorldQuant BRAIN* operates as a simulation platform that institutionalizes external quant research, further proving that algorithmic prediction is evolving into a distributed, competitive ecosystem rather than a localized institutional secret (WorldQuant, 2025).

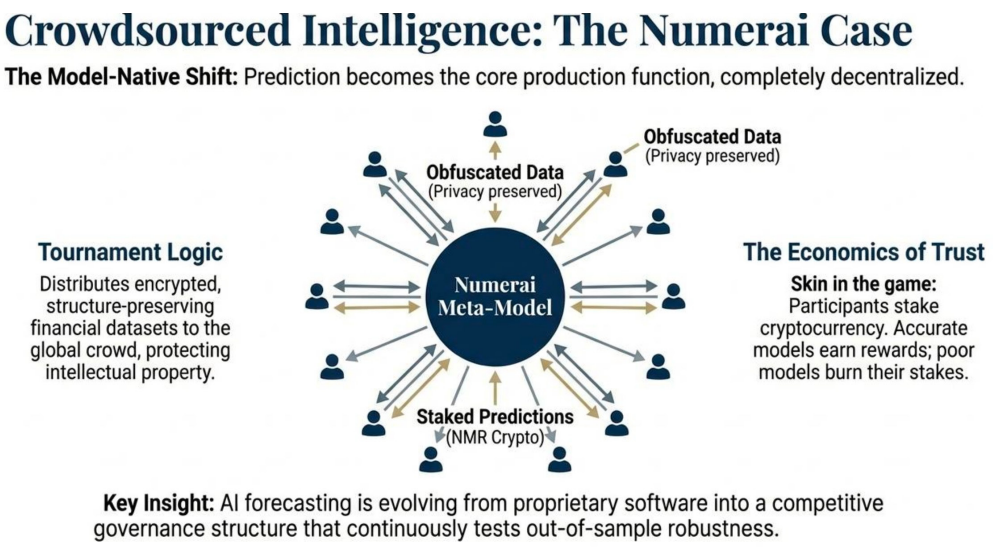


Figure 3: Crowdsourced Intelligence

3.6 Reconfiguring human expertise: The hybrid profile

Ultimately, the empirical evidence from both platform implementation and AI-native models demonstrates that AI does not replace human experience; it reconfigures it. The traditional investment banker is evolving into a "hybrid profile". Analysts who delegate computational forecasting to models do not abdicate professional judgment. Instead, their cognitive efforts shift *upstream*—to problem definition, feature engineering, and rigorous data governance—and *downstream*—to robustness testing, exception management, and controlling for algorithmic drift (Lopez de Prado, 2020). The integration of conversational modules, such as Aladdin-Copilot, further supports this by translating complex risk scenarios into natural language, assisting the "human in the middle" in preparing client narratives and ensuring fiduciary suitability before executing trades (BlackRock, 2024).

4 Discussion

The findings of this study necessitate a critical discussion regarding the socio-technical limitations, systemic risks, and governance imperatives surrounding the integration of AI in investment management. The analysis is structured around the conceptual framework of structural limits and strategic synthesis.

4.1 Structural limits and technical failures associated with financial AI

While the methodological shift toward non-linear machine learning expands predictive boundaries, the discussion must emphasize the inherent technical failures associated with financial AI. Financial markets are characterized by non-stationarity and reflexivity; the statistical properties of asset returns shift continuously due to changing macroeconomic policies, participant behaviors, and structural breaks (Lo, 2004; Potters, Cont, & Bouchaud, 1998). Consequently, predictive models are highly susceptible to *overfitting*—the tendency of an algorithm to capture historical noise rather than economically meaningful signals, resulting in predictive fragility when applied to out-of-sample, real-world data. This fragility was starkly evident during the COVID-19 pandemic shock of March 2020, where models trained exclusively in low-volatility regimes failed dramatically. Furthermore, the pervasive risk of *data leakage*—where future information inadvertently contaminates training datasets—can create a dangerous illusion of algorithmic precision, misleading risk management committees (Lopez de Prado, 2020). The inherent limits of AI dictate that while computational power has scaled exponentially, the signal-to-noise ratio in financial forecasting remains stubbornly low.

4.2 Market dynamics and systemic risks

From a macro-institutional perspective, the widespread adoption of AI introduces severe systemic risks tied to market dynamics. Financial markets are adaptive ecosystems. When multiple institutions deploy similar AI architectures, consume the same alternative data feeds, and utilize homogenized platforms like Aladdin, the sector faces a high risk of algorithmic convergence (Financial Times, 2020). This convergence implies that market participants may react uniformly and simultaneously to specific algorithmic stimuli, potentially causing "herd behavior" that exacerbates market selloffs and amplifies systemic volatility, reminiscent of the 2010 Flash Crash (BIS, 2024; FSB, 2024). The democratization of platforms like VantagePoint and Tickeron across retail spaces further compound this issue; if a collective mass of retail algorithms executes simultaneous breakout strategies, they can create self-reinforcing price shocks. Furthermore, technological dependence on a concentrated group of dominant cloud and AI providers exposes the global financial system to amplified operational and cyber risks (Zetzsche, Buckley, & Arner, 2020).

4.3 Institutional concerns and algorithmic governance

The tension between maximizing predictive accuracy (often via complex deep learning networks) and maintaining interpretability creates profound institutional concerns. In the heavily regulated domain of investment banking, relying on "black box" algorithms poses severe fiduciary and legal risks (Goodfellow, Bengio, & Courville, 2016). If an AI system systematically denies credit, misallocates capital, or executes a highly detrimental trade, institutions struggle to map the precise algorithmic logic to assign legal accountability (Goodman & Flaxman, 2017). Data bias poses an additional threat; algorithms trained on historically prejudiced datasets can systematically discriminate, as observed in consumer lending anomalies (Bartlett et al., 2019; O'Neil, 2016; Mehrabi et al., 2021). While post hoc Explainable AI (XAI) techniques offer partial transparency by highlighting feature importance, they do not resolve the fundamental opacity of deep neural networks (Barredo Arrieta et al., 2020; Quinn, 2023).

This opacity is directly confronted by the regulatory landscape, most notably the European Union's AI Act (European Commission, 2021; European Commission, 2023). The AI Act acts as a catalyst for algorithmic governance by introducing a risk-based framework in which certain financial AI applications—particularly those affecting credit scoring, risk assessment, or other high-impact decision-making processes—may fall within the high-risk category. For such systems, the regulation imposes stringent requirements for data quality, human oversight, traceability, and rigorous conformity assessments. Consequently, the discourse surrounding AI is forced to transition from a narrative of pure operational efficiency to an imperative of corporate ethics and governance. The deployment of AI necessitates a "new responsibility pact," meaning that the promise of rapid, highly personalized, data-driven financial advice must be inexorably linked to guarantees of fair, defensible, non-discriminatory, and highly verifiable decision-making structures.

4.4 Strategic synthesis: human-in-the-loop and financial Judgment

Synthesizing these technical capacities and structural risks reveals a new strategic reality for investment banking. AI cannot eliminate the ontological uncertainty of markets; instead, it serves to provide *probabilistic*

decision support. The competitive edge for modern institutions is no longer derived from discovering a flawless, deterministic oracle, but rather from securing *marginal accuracy*—a slight probabilistic tilt toward outperformance that, when applied across multi-trillion-dollar portfolios under strict risk constraints, yields massive economic value.

Achieving this requires a fundamental *transformation of financial judgment*. Judgment is no longer the sole purview of the isolated star analyst; it is the product of an ensemble of models governed by strict institutional protocols. The most resilient organizational architecture is one grounded in *human-in-the-loop governance*. Human professionals are not rendered obsolete; they are elevated to the role of systemic overseers. While AI handles high-dimensional signal extraction and pattern recognition, human actors provide the critical contextual interpretation of novel geopolitical events, manage ethical compliance, and bear the ultimate fiduciary responsibility that algorithms cannot legally assume.

5 Conclusions, Limitations, and Future Research

This study concludes that Artificial Intelligence in investment banking constitutes a structural reconfiguration of the financial architecture rather than a mere technological optimization. AI does not act as a deterministic oracle capable of guaranteeing return forecasts; its foundational value lies in providing disciplined, probabilistic advantages under conditions of deep uncertainty. Digital investment platforms, functioning as cognitive infrastructures, successfully facilitate human-AI synergy by synthesizing vast data layers, enabling complex non-linear machine learning, and standardizing operational processes. However, this synergy is inherently precarious. Sustainable innovation within the financial sector requires these cognitive infrastructures to be firmly embedded within robust algorithmic governance frameworks, particularly in light of stringent mandates such as the EU AI Act (European Commission, 2023; OECD, 2023). Intelligent finance is only viable if the algorithmic systems driving it are transparent, auditable, and ethically grounded by human professionals who maintain ultimate fiduciary authority.

The research acknowledges several methodological limitations. The qualitative nature of the empirical investigation, based on semi-structured interviews concentrated within the European Union banking sector, restricts the broader global generalizability of the findings. The analysis lacks access to proprietary quantitative datasets and the underlying source code of the platforms examined, relying instead on user perspectives, public documentation, and literature. Furthermore, the study observes the regulatory environment during the early, highly dynamic implementation phase of the EU AI Act, meaning that long-term compliance impacts are yet to be fully realized.

Future research must expand upon these findings by investigating the longitudinal effectiveness of Explainable AI (XAI) in meeting evolving regulatory compliance standards across divergent global jurisdictions (e.g., comparing EU, US, and Asian regulatory environments). Subsequent studies should critically evaluate the political economy of platform-based financial infrastructures, specifically quantifying the systemic implications of technological dependence and the concentration risks posed by algorithmic convergence (Zetzsche, Buckley, & Arner, 2020). Finally, empirical research is needed to track the long-term economic performance and behavioral stability of retail AI forecasting tools and crowdsourced AI-native ecosystems during periods of extreme macroeconomic shock.

6 Ethics Declaration

Ethical clearance was not required for the research

7 AI Declaration

AI tools were used for revising English language

References

Arner, D. W., Barberis, J., & Buckley, R. P. (2016). The evolution of Fintech: A new post-crisis paradigm? *Georgetown Journal of International Law*, 47(4), 1271–1319.

Barredo Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., & Herrera, F. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, 58, 82–115.

Bartlett, R., Morse, A., Stanton, R., & Wallace, N. (2019). Consumer-lending discrimination in the FinTech era. *Journal of Financial Economics*, 143(1), 30–56.

BIS – Bank for International Settlements. (2024). Artificial Intelligence and the Financial System, Basel: BIS; Annual Economic Report 2024: Chapter III (Artificial intelligence and the economy), Basel: BIS.

Brynjolfsson, E., & McAfee, A. (2017). *Machine, platform, crowd: Harnessing our digital future.* New York: W.W. Norton & Company.

Buehlmaier, M., & Whited, T. (2023). Artificial intelligence in asset management: Evidence from sentiment analysis. *Journal of Finance*, 78(5), 2417–2459.

Cao, L. (2024). *AI in Finance: Applications, Challenges, and Emerging Trends*. Springer.

Chen, C., Härdle, W. K., & Ong, B. (2021). AI in finance: Theory and applications. *Journal of Financial Data Science*, 3(2), 1–12.

Davenport, T. H., Guha, A., Grewal, D., & Bressgott, T. (2020). How artificial intelligence will change the future of marketing. *Journal of the Academy of Marketing Science*, 48(1), 24–42.

European Commission. (2021). Proposal for a Regulation laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act), Brussels.

European Commission. (2023). Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act), Brussels.

Feng, G., He, J., & Polson, N. (2018). Deep learning for predicting asset returns. *Applied Stochastic Models in Business and Industry*, 34(1), 1–12.

Financial Times. (2020, March 1). How BlackRock's Aladdin became the technology behind Wall Street. *Financial Times*. Retrieved September 30, 2025.

FSB – Financial Stability Board. (2024). Artificial Intelligence and Machine Learning in Financial Services. Basel: FSB.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.

Goodman, B., & Flaxman, S. (2017). European Union regulations on algorithmic decision-making and a "right to explanation". AI magazine, 38(3), 50-57.

Heaton, J.B., Polson, N.G., Witte, J.H. (2016). *Deep learning for finance: deep portfolios*. Wiley Online Library.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2), 263–291.

López de Prado, M. (2020). *Machine learning for asset managers*. Cambridge University Press.

Malo, P., Sinha, A., Takala, P., Korhonen, P., & Wallenius, J. (2014). Good debt or bad debt: Detecting semantic orientations in economic texts. *Journal of the Association for Information Science and Technology*, 65(4), 782–796.

McKinsey & Company. (2023). *The state of AI in financial services 2023*. McKinsey Global Institute.

Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A survey on bias and fairness in machine learning. *ACM Computing Surveys*, 54(6), 1–35.

Nakamura, T. (2022). Machine learning applications in financial forecasting: From linear models to deep learning. *Journal of Economic Perspectives*, 36(2), 45–68.

O'Neil, C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and threatens democracy*. New York: Crown Publishing Group.

OECD. (2023). AI, Ethics and Financial Regulation: Building Trustworthy Systems in Finance. Paris: OECD Publishing.

Potters, M., Cont, R., & Bouchaud, J. P. (1998). Financial markets as adaptive systems. *EPL* (Europhysics Letters), 41(3), 239-244.

Quinn, J. (2023). AI explainability in financial markets: Balancing transparency and performance. *Financial Innovation Journal*, 9(12), 1–23.

Ruff, L., Vandermeulen, R., Goernitz, N., Deecke, L., Siddiqui, S. A., Binder, A., & Kloft, M. (2021). Deep one-class classification. *Proceedings of the 36th International Conference on Machine Learning* (ICML), 97, 4393–4402.

Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *Journal of Finance*, 62(3), 1139–1168.

Varshney, M. M., & Varshney, P. K. (2026). "AI and Fintech synergy: Strengthening financial stability in Islamic and conventional banks", *Journal of Risk and Financial Management*, *19*(1), 21.

Zetzsche, D. A., Buckley, R. P., & Arner, D. W. (2020). Artificial intelligence in finance: Putting the rules in place. *Journal of Banking Regulation*, 21(3), 283–297.

Platforms sitography

BlackRock. (2024). Aladdin: The investment operating system. BlackRock. Retrieved September 30, 2025.

BlackRock (2024), Aladdin Enterprise, Risk, Data Cloud and Accounting, BlackRock Aladdin official pages, Retrieved September 30, 2025

Danelfin, AI Score methodology & stock analytics platform, Retrieved from https://danelfin.com/, 2025

Numerai, Numerai platform overview, Retrieved from https://numer.ai, 2026

Numerai Docs, Overview & staking documentation, Retrieved from Tickeron, AI Tools for stock forecasts & price predictions, Retrieved from https://tickeron.com/, 2025

VantagePoint AI, Artificial Intelligence Trading Software – market forecasting, Retrieved from https://www.vantagepointsoftware.com/, 2025

Reuters, AI drives surge in WorldQuant's university quant contest participation, Retrieved from https://www.reuters.com/technology/ai-drives-surge-worldquants-university-quant-contest-participation-2025-10-06, 2025

Annex 1 Questionnaire

Introductory / Context-setting

- What's your name and which role do you play within your organisation?
- Could you tell me about your current role and how long you've been using the Aladdin platform?
- How was Aladdin introduced within your organisation? What were the needs?
- Before adopting Aladdin, what kind of tools or processes were you using for portfolio and risk management?

Operational Use and Integration

- In which phases of your work do you mainly use Aladdin — portfolio construction, risk monitoring, reporting, or client advisory?
- How are you training or updating to make the best use of the platform and fully exploit its operational potential?
- How does Aladdin interact with other internal systems (CRM, accounting, ESG databases, etc.)?
- From your experience, what are the main benefits in terms of speed, efficiency, or data consistency compared with legacy tools?
- To what extent do you rely on Aladdin's outputs versus your own professional judgment when taking investment decisions?

Artificial Intelligence and Analytics Capabilities

- Which AI- or machine-learning-based functions within Aladdin are most relevant for your activity?
- Do you perceive these tools as decision-support instruments or as autonomous decision-makers?
- How transparent are the algorithms' results to you — do you understand why a certain recommendation or risk alert is generated?
- How does your team validate or challenge model outputs before acting on them?

Data Quality and Governance

- How is data quality managed within the platform? Do you encounter issues of completeness, timeliness, or bias in the datasets used?
- Are there internal procedures or controls to ensure that Aladdin's data inputs are consistent with your bank's risk-governance standards?
- In your opinion, how critical is data governance for achieving reliable AI-driven insights?

Human Factor and Professional Culture

- Has the adoption of Aladdin changed the required skill-set for professionals in your division?
- How do you balance quantitative model outputs with qualitative considerations, such as client expectations or market intuition?
- Do you think younger generations of analysts relate differently to Aladdin compared with senior portfolio managers?

Client Relationship and Transparency

- How do you explain to clients that part of the investment process is supported by AI-driven analytics?
- Have clients ever expressed concerns about automation, opacity, or loss of human oversight?
- Does Aladdin help you enhance the personalisation of advisory services (e.g., ESG preferences, behavioural analysis)?

Risk Management and Regulatory Compliance

- How does Aladdin assist you in identifying and monitoring portfolio risks such as VaR, stress testing, or concentration risk?
- In what ways does the platform help fulfil regulatory requirements (e.g., MiFID II, ESG disclosures, internal audits)?
- How is accountability managed if an AI-driven recommendation leads to a loss — who bears responsibility?

Strategic Vision and Future Outlook

- In your view, what is the next evolutionary step for platforms like Aladdin — more automation or more human interpretability?
- How do you imagine the ideal interaction between human advisors and AI systems five years from now?
- What advice would you give to a financial institution that is about to implement Aladdin for the first time?

Digital Innovations in Sustainable Supply Chain Business Models

Maria Rosaria Marcone
Marche Polytechnic University, Ancona, Italy

Abstract

The aim of this work is to identify the new innovative phenomena affecting Supply Chain Business Model Systems (SCBMSs), with reference to supply chain relationships that become digital. The focus is on upstream supply chain operators and leading supply chain firms.

We hypothesize that leading firms (almost always small and medium-sized) providing these technologies need to search for external knowledge, improve cooperation, develop servitized offers and connectivity along the value chain. These approaches allow wider customer contact, easier communication, and automated value-capturing and value co-creation processes. Supply chain platform development, and servitization are important in providing advanced digital technologies solutions. We focus on the notion of knowledge complexity, which reflects both the variety and newly-rarity of the knowledge stock available in a given supply cain ecosystem. Existing research on the relationship between strategic renewal and corporate purpose appears fragmented and provides mixed insights. Some evidence points to purpose-driven firms as being more flexible in navigating change better able to radically reinvent themselves. Regarding the methodology, this research used theoretical sampling for case selection since the objective is to build concepts and theory regarding the mechanisms involved in supply chain business model (SCBM) when digital technologies challenge factors are present. Therefore, the main criterion for sample definition was to search for cases representative of I4.0 technology provision regarding key technologies offered with leading representativeness of the technologies and companies analysed in the manufacturing supply chain (SC). Our results show that the I4.0 technologies and solutions impact the whole SCBMSs. This paper can contribute to both theory building and provide important practical implications by: (a) enriching the conceptual understanding of the unintended consequences of digital transformation; (b) providing managers important insights and practical advice for riding on the digitalisation to enhance operations-supply side relationships and supply chain value offered.

Keywords: digital supply chain, business model innovation, ecosystems, digital supply chain platforms.

Paper type: Academic Research Paper

1 Introduction

The aim of this work is to identify the new innovative phenomena affecting Supply Chain Business Model Systems (SCBMSs), with reference to supply chain relationships that become digital (Müller et al., 2021; Marcon et al., 2022). The focus is on upstream supply chain operators and leading firms (sometimes their business-industrial clients); the supply chains studied are truly innovative ones such as the machinery, biomedical, agri-tech food, chemical, medtech supply chains.

We hypothesize that SC leading firms (almost always small and medium-sized) providing these technologies need to search for external knowledge, improve cooperation, develop servitized offers and connectivity along the value chain (Frank et al., 2022; Besiroglu et al., 2024; Grimm et al., 20242). These approaches allow wider customer contact, easier communication, and automated value-capturing and value co-creation processes.

Existing research on the relationship between strategic renewal and corporate purpose appears fragmented and provides mixed insights. Some evidence points to purpose-driven firms as being more flexible in navigating change better able to radically reinvent themselves.

The field of this study is digital transformation both in firm's sustainable strategies and in SC restructuring or redesign. Digital transformation is a process of integrating digital technologies into all firms and SC business activities: a mega-process that involves substantial changes in technology, business management culture, design and engineering operations in production systems, and supply chain value generation (macro-level changes) (Thomas and Ritala, 2022; Teoh et al., 2023; Laamanen et al., 2025). The research perspective is the most recent one of innovation that considers the prerogative of digital servitization which favours the creation of a new serviced business model by implementing digital resources for the creation of new sustainable – at firm level and for the complete SC - value. The comprehensive development of digitalising sustainability information for communication across supply chain actors has become increasingly important.

The research's focuses are as follows. The first is about the SC digital innovative sustainability evolution. Supply chain's firms that attach strategic importance to environmental issues are those that recognize a new opportunity to compete. They are inclined to make large investments intended for the adoption of non-polluting

technological solutions in production systems, as well as to initiate internal training activities to increase human resources' environmental skills and also to design specific and well-detailed internal managerial procedures. Sustainability is a potential competitive advantage. Sustainable development is currently a pivotal element in the processes of formulating the competitive strategies of many enterprises (we will deal with manufacturing ones) and presupposes that additional objectives compared to traditional ones (economic-financial objectives) are considered both in the formulation of growth strategies and in the design of modern internal governance procedures.

The second is the renovation or even design of new Supply Chain Business Model Systems (SCBMSs).

The business model of modern supply chains must necessarily be restructured by the companies operating there, by new entrants, and more generally by economic innovation actors, including those external to the SC, who, together with SC companies, are included in new open innovation or open innovation (OI) systems. All these economic innovation actors are oriented towards capturing early signs of change coming from outside (digital technologies, new firms as the start-up, institutional incentives and regulatory laws, etc.) or even acting proactively on the external context (supply relationships, the relevant social environment, new and more advanced digital strategies than in the past.

Business model modification is concretely and profitably possible in the production structures and governance structures of the SCs to which they belong, which with agility (operational agility) adapt to new and disruptive technologies and new strategic drives through which to compete (green innovation, sustainability, digital SC relationships up stream and down stream).

This research explores DT technology integration into global manufacturing firms from an operational capability perspective. The research followed the single case study guidelines and took additional inspiration from longitudinal case study guidelines and process study guidelines. Several sources for data analysis are available for case study research. This helped highlight overlapping or emerging concepts, themes, and dimensions as a contribution of new theoretical insights.

2 Theoretical Framework

The modification of the SCBM and the new value creation resulting from it are sources of a competitive advantage that has also been modified, as it differs, sometimes significantly, from the past. Moreover, in strategic management studies, traditionally focused on decision formulation, increasing importance is recognized in the analysis of how changes are generated and in the evaluation of the effects they generate (Frankenberger and Sauer, 2019).

The production of scientific contributions on the topic of business models is more active and livelier than ever. However, there is a lack of references and above all field explorations on the birth of new SCBMs. Some useful insights can start from the reflections that some authors have started on topics, reported below: Brucherer (2012) recognizes two important causes of modification of business models 'macro-dimensions', the market and industry; Teece (2010) argues that the criticality of the operational meaning of the business model lies in the connection it makes possible between the 'micro-foundations' or 'micro-elements' constituting the business strategy and the competitive performance actually achieved by it; other authors (Blackburne and Buckley, 2019) are among the most convinced supporters of the business model as an operational process in continuous evolution; others state that supply chain business model innovation (SCBMI) mediates the link between technological innovation and supply chain performance (Tate et al., 2010; Karimi-Walter, 2016; Gölgeci et al., 2018). These authors define the SCBM is defined as new, when it is the result of the design, by managers and entrepreneurs, of a radically new system of activities for objectives, structure and forms of governance (architectural novelty of the business) or a system, in which activities with elements of novelty (objectives, structure, governance) integrate with unmodified activities (complementary novelty of the business).

Knowledge management (KM), its growth or disruptive development is an important research field and the basis for the performance of effective managerial procedures and more than ever open to new digital technologies.

The knowledge management in innovative and sustainability-oriented supply chain management procedures includes key firm's areas of enterprise and business partner's areas with which it interacts.

KM in SCM that includes many kinds of business partner interaction, joint knowledge creation and knowledge dissemination at both inter- and intra-organizational levels to enhance the efficiency of supply chain activities (Pahm et al., 2024). KM in SCM uses a range of advanced technologies for capturing and processing supply chain-related information. This acquired and refined information (i.e. knowledge) is circulated within firms and their externalities to make timely decisions and thereby sustain competitiveness (Massa et al., 2023; Gustafsson et al., 2024).

RQ 1. KM strategies generate (this binominial) structural managerial procedure changes of SCs.

Medium sized firms (MSFs) employ OI practices (Livieratos et al. 2024) to overcome the challenges associated with successfully utilizing digital innovative technologies (Hongyun et al. 2023; Meng et al. 2021). This research gap raises an important research question about what the challenges in SMEs are leveraging business model and related digital innovations and what particular OI practices through collaborations and partnerships with external innovators SMEs employ to overcome the challenges.

Indeed, MSFs with resource constraints may trigger different OI practices based on their situational experience (Spithoven et al. 2013; Brunswicker and Vanhaverbeke 2015; Massa et al. 2023). Our study explores (1) the challenges in MSFs leveraging business model and related digital innovations for service delivery; (2) the specific OI practices through collaborations and partnerships with external organizations employed by SMEs; and (3) proposes a framework that connects OI collaborations and partnerships to the specific challenges faced by SMEs leveraging business model and related digital innovations. In modern digitalized and sustainability-oriented SCs, innovations are inspired by principles never experienced in the past. Attention to open innovation may appear as a strategy for firms to respond to technological adoption, there is limited research on. We investigate as follows.

RQ 2. OI is a great opportunity for improve firm's knowledge management.

Digital supply chain platforms are emerging as critical infrastructures that empower the digital transformation of manufacturing, attracting growing academic and practical interest. By leveraging modular digital architectures, IDPs enable extensive interconnectivity among humans, machines, physical objects, and systems, thereby facilitating the integration and sharing of heterogeneous resources (Pauli et al., 2025; Bai et al., 2026).

Also, the platform-native firms, as cross-boundary entrants, orchestrate value co-creation strategies—resource integration, service exchange and institutional arrangements—to develop digital platforms in new and modern SC. We believe, welcoming the analyses of Qin and Li Sun (2026) that the ability to combine knowledge is an important mediator. Additionally, we introduce the moderating role of the ease of knowledge combination, investigating how it interacts with knowledge dependence to influence innovation outcomes. Specifically, how it facilitates the adoption of sustainable technologies in increasing opportunities to continuously improve capacities to successfully introduce digital technologies. We adopt the concept of knowledge dependence to explain how suppliers drive innovation and capture value within value chains. Knowledge dependence among SC's levels is central to this process, as the need to obtain external and complementary knowledge promotes interconnections and interdependence among firms (Kristiansen, 2026)

RQ 3a. The platform generates supply chains restructuring.

Regarding the s it has been studied how the digital technologies facilitate diverse functions within digital servitization, including monitoring, controlling, automating processes, and creating fully autonomous solutions (Kohtamäki *et al.*, 2019; Kristiansen, and Aas 2026). Other studies have highlighted that successful digital servitization relies on technical capabilities and the willingness of ecosystem actors to cooperate, particularly concerning data exchange (Kolagar et al., 2022). In addition to technical factors, social elements must be considered when assessing the feasibility digital servitization outcomes (Dalenogare et al., 2023). Non-large enterprises can also be leading players (Khan et al., 2026). the digital platform generates the servitization of the supply chain (Orgink et al., 2023).

Supply chain platform development, and servitization are important in providing advanced digital technologies solutions. We focus on the notion of knowledge complexity, which reflects both the variety and newly-rarity of the knowledge stock available in a given supply cain ecosystem (Hillebrand et al., 2025). Variety as increasing complexity that has positive impact on recombinant knowledge dynamics as it supports the scope for the recombinant generation of new knowledge and eventually leads to augmented rates of introduction of innovations.

This paper also sheds light on how innovative resources interact with other resources to drive servitization, bridging the gap in research between servitization and innovation (Coreynen et al., 2017; Kohtamäki et al., 2019).

RQ 3b. Digital SC platform as digital SC servitization.

3 Methodology

This research uses theoretical sampling for case selection since the objective is to build concepts and theory regarding the mechanisms involved in supply chain business model (SCBM) when digital technologies challenge factors are present (Thomas and Ritala, 2022). Therefore, the main criterion for sample definition was to search for cases representative of I4.0 technology provision regarding key technologies offered with leading representativeness of the technologies and companies analyzed in the manufacturing industry.

In the second step, we analyzed the technologies offered by these companies to consider a diversity of technologies related to I4.0 solutions and the profile of the companies' business models, i.e., whether the business models were strongly oriented to I4.0 solutions and not only the provision of single technologies. In this sense, although some companies focus on specific technologies, all selected provide integrative solutions to manufacturers based on the leading technology.

Qualitative research has been chosen. The research framework design emerges through an inductive analysis method with the intention of having a positive-interpretative position, in order to highlight emerging innovative phenomena (Figure 1).

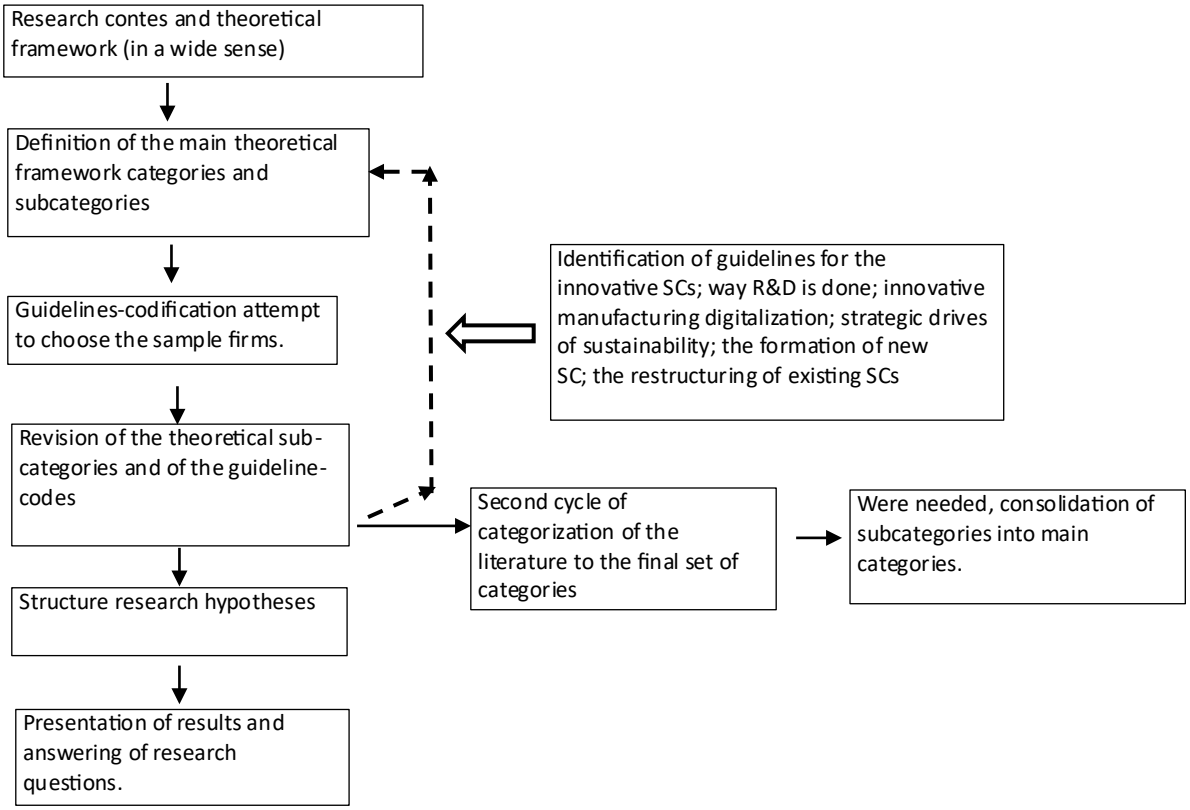


Figure 1: Steps of research process
Source: our elaboration

An interpretative, qualitative approach utilizing selected multi-case study analysis is chosen: drawing on qualitative analysis, it was first investigated how companies belonging to their supply chain design their research activities, to generate sustainable innovative processes (sheet 1), and then the reconfiguration of sustainable innovative supply chain, sustainable strategic choices of supply chain companies, strategies and structural choices (way of doing research, redesign of supply side relations, digitization of the supply chain relationships, foreign development). Several sources for data collection are available for case study research.

We used an explorative approach, conducting semi-structured interviews with firms' managers, sustainability experts, researchers, etc. Analysis of sector and sub-sectors data based has been conducted on the re-elaboration of industry associations reports (both national and European reference) (see sheet 1). Data were collected from 2024 to these first months 2026 through face-to-face interviews conducted at participants' offices or neutral meeting locations in Italy. Conducting interviews in person allowed for richer interactions, capturing non-verbal cues and fostering a deeper rapport with participants. The focus on a single country ensures consistency in cultural, regulatory, and economic contexts, which are critical factors influencing the adoption of AI technologies.

The aim of the work is to follow the strategic development of the cases over time (evolution, growth, etc.). The iterative analysis which perhaps better than other approaches, it is believed, may prelude a longitudinal analysis. Instead of testing a predefined hypothesis, this study aims to generate insights through an open-ended, iterative process for explorative data collection from cases. This study approach is technically possible for the supply chain leading firms; not applicable for startups, spin-offs, institutional stakeholders, because the first two are too young and are continuously subject to change, with regard to content, institutional interventions.

Sheet 1: Methodology choices

A) Strategic drives

The emphasis on the SCBMI has led to the direction of field experimentation, privileging, in this step of empirical research which is still being completed, the investigation of the development strategic drives which are expressed in the table that follows (see table 1).

Table 1: Strategic drives

Strategic drives	SC micro-foundation practices
(1) SC value co-creation	- attitudine e l'impegno ai progetti collaborativi di miglioramento - estensione delle interdipendenze fra le imprese all'interno della SC (relativa importanza di un cliente per il business del fornitore e criticità delle forniture) - problem solving
(2) Ways of doing innovation	- OI - innovation community - other external drive
(3) Relationship efficiency	- the duration of the relationship and the types of past experiences; - technological or procedural links between firms and among diversified SCs;
(4) Digital relationships	platformization (digital relationships and role of the leading firms)
(5) Governance	the length and complexity of the Supply Chain

B) Source data collection

Primary data were collected as shown in the following: (a) Semi-structured interviews in trade fair forums; (b) Industry conferences; (c) Semi-structured Interviews in meeting-round table; (d) Semi-structured Interviews in visiting (CPO, CFO); (e) Semi-structured Interviews in scientific Italian managerial conference (CEO, CPO, CFO); (f) Direct observations.

Secondary data were collected as shown in the following: (1) Press coverage, company blogs, analyst reports; (2) Interviews given by the area manager and published in European media outlets; (3) Reports of sector research centres and consultancies; (4) Company blogs, analyst reports (Make Up the future)-2026. Champions-Italy Post Ranking.

C) R&D activities

Strategic choices were analysed, favouring the analysis of new and unusual ways (due to the size of economic actors and companies and the sectors from which they originate) of conducting R&D within the SCs of origin and in the new and diversified ones that necessarily arise (see table 2)..

Table 2: R&D activities

Cases	Innovative activities in sustainable supply chains
Diary	R&D in the various supply chain levels Diversification: multi-vertical supply chain integration
Medtech	Joint R&D of cooperatives in research HUB R&D in platformed supply chain.
Agrochemical	Internal R&D shared with suppliers (Supply side integration) OI in know how increasing in partnership Design and engineering of production processes
Agrifoodtech	Research hub R&D and design 'Hub Bite'
Advanced mechanics	Open Innovation Platform with innovative SC: solutions for next generation mobility
Agritech	New R&D in collaboration with cluster partners (Ferrara-Italy) operating in machinery, animal husbandry and poultry farming. New R&D in collaboration with a partner of Milano belonging at the sub-sector 'food supplements'.
Coffetech	AI to increase soil fertility (new plantings on region affected by floods, desertification, fire)
Digital cosmetics (two cases)	NPD for saving scarce resources (water) Radical innovation raw materials AI for data processing and quality control (traceability, digital twin in new product design-engineering-production)

4 Conclusions

The main challenge in comprehending the unintended consequences of digital transformation is our existing knowledge base. Digital transformation and the shift toward sustainable innovative strategies at the level both the firms and the SCs are reshaping many supply chains, raising both practical and theoretical questions about power, collaboration, and ecosystem dynamics. Considering this, interest in supply chain digitalisation have grown significantly in recent years.

Our study makes two primary contributions. First, we extend OI research in no large firms by illuminating how the challenges faced by analysed firms when leveraging digital innovation. Second, our work provides valuable empirical support for OI target strategies understanding (RQ 2). These strategies are not only relevant to no large firms but also extend the interest in them by those businesses and economic and innovation actors that to other firms with limited financial resources and moreover interested in expanding (diversifying) the portfolio of businesses, even diversified SCs, to respond to and even improve growth opportunities arising from the disruptive advent of digital sustainability. We find that among the startups participating in the formation of new production chains, it emerges that those that achieve better performance from the first years of life are those that come from 'contexts' (industrial districts, sustainable innovation ecosystems, research and innovation hubs) where events of sedimentation and growth of knowledge management have occurred in the past. It is as if the localization of knowledge has its own differentiated 'propagation' of innovation: from localized specificities further localized innovative specificities arise which then internationalize and therefore become multi-localized, even if not in a merely replicative sense (traditional model of internationalization) (RQ 1).

The study finds and observes that suppliers-manufacturer relationships should first and then prioritize investments in new digital technologies for the short, medium, and long term regardless of the evaluation of their current operational capabilities. This is the emergence field of the digital supply chain platforms thus finding RQ 3a. Supply chain digital platforms enable the real-time interconnection and sharing of dispersed physical manufacturing resources, data resources, and knowledge resources among multiple actors. This facilitates efficient cyber-physical integration, establishing the foundation for subsequent collaborative development of smart manufacturing service solutions. Digital platforms enable firms to overcome classical information and coordination constraints. Unlike traditional intermediaries (such as brick-and-mortar travel agencies), digital platforms enable firms to overcome classical information and coordination constraints by combining market mechanisms with elements of hierarchical control in hybrid, firm-designed systems. These platforms rely on semi-directed and generative search processes, which go beyond centralized mechanisms to enhance innovation and adaptability in fast-changing environments. In addition, platforms provide a low-cost pathway to scale, offering superior capabilities for environmental scanning, opportunity sensing, and the integration of diverse inputs. This paper also sheds light on how innovative resources interact with other resources to drive servitization, bridging the gap in research between servitization and innovation (RQ 3b). Digital transformation of business models is a SC digital servitization that improve - aiming to revamp and reinvigorate - their operations to better engage, and create new values for their clients, suppliers and communities. The result is proactivity. The advantages of digital transformation are enormous as pointed out by many authors. Digitalisation as transformative servitization interests the processes of the SCBM: they adopt digital technologies to change operations and automate processes. The result is reactivity.

Supply chain platform development, and servitization are important in providing advanced digital technologies solutions. We focus on the notion of knowledge complexity, which reflects both the variety and newly-rarity of the knowledge stock available in a given supply cain ecosystem (RQ 1). Variety as increasing complexity that has positive impact on recombinant knowledge dynamics as it supports the scope for the recombinant generation of new knowledge and eventually leads to augmented rates of introduction of innovations.

This research provides a significant contribution to knowledge to digitalisation, digital transformation, and supply chain relationship related research. Our findings extend existing literature by identifying underexplored unintended consequences of digital transformation in circular economy contexts, particularly in relation to social equity, digital inclusion. This paper can contribute to both theory building and provide important practical implications by: (a) enriching the conceptual understanding of the unintended consequences of digital transformation; (b) providing managers important insights and practical advice for riding on the digitalisation to enhance operations-supply side relationships and supply chain performance. Our results show that the I4.0 technologies and solutions impact the whole SCBMSs.

5 Ethics Declaration

It is stated that for the research presented in this paper, it was not necessary to apply for an ethical authorization.

6 AI Declaration

It is declared that no artificial intelligence tools were used.

Referencs

Bai, W., Wu, D., Yang, M. and Tanfei Liu, T. (2026) "Value co-creation in B2B industrial digital platforms: A platform-native perspective under service-dominant logic", *Technovation*, Vol. 153, 103516,

Besiroglu, T., Emery-Xu, N. and Thompson, N. (2024) "Economic impacts of AI-augmented R&D", *Research Policy*, Vol. 53, No. 7, 105037.

Blackburne, G. D. and Buckley, P. J. (2019) "The international business incubator as a foreign market entry mode", *Long Range Planning*, Vol. 52, No. 1, pp. 32-50

Brucherer, E., Eiser, U. and Gassman, O. (2012) "Towards systematic Business Model Innovation: Lessons from Product Innovation Management", *Creativity and Innovation Management*, Vol. 21, No. 2, pp. 183-198.

Brunswicker, S. and Vanhaverbeke, W. (2015) "Innovazione aperta nelle piccole e medie imprese (PMI): strategie di acquisizione di conoscenza esterna e facilitatori organizzativi interni", *Journal of small business management*, Vol. 53, N. 4, pp. 1241-1263.

Coreynen, W., Matthyssens, P. and Van Bockhaven, W. (2017) "Boosting servitization through digitization: pathways and dynamic resource configurations for manufacturers", *Industrial Marketing Management*, Vol. 60, pp. 42-53.

Dalenogare, L.S., Le Dain, M.-A., Ayala, N.F., Pezzotta, G. and Frank, A.G. (2023) "Building digital servitization ecosystems: an analysis of inter-firm collaboration types and social exchange mechanisms among actors", *Technovation*, Vol. 124, 102756.

Frank, A. G., de Souza Mendes, G. H., Benitez, G. B. and Ayala, N. F. (2022) "Service customization in turbulent environments: Service business models and knowledge integration to create capability-based switching costs", *Industrial Marketing Management*, Vol. 100, pp. 1-18.

Frankenberger K. and Sauer R. (2019) "Cognitive antecedents of business models: Exploring the link between attention and business model design over time", *Long Range Planning*, Vol. 52, No. 3, pp. 283-304.

Gölgeci, I., Murphy, W. H. and Johnston, D. A. (2018). "Power-based behaviors in supply chains and their effects on relational satisfaction: A fresh perspective and directions for research.", *European Management Journal*, Vol. 36, No. 3, pp. 278-287.

Grimm, J., Langley, A. and Reinecke, J. (2024) "Process Research Methods for Studying Supply Chains and Their Management", *Journal of Supply Chain Management*, Vol. 60, No 4, pp. 3-26.

Gustafsson, C., Chirumalla, K., Bruch, J. and Sannö, A. (2025) "Additive manufacturing technology integration: an operational capability perspective", *Journal of Manufacturing Technology Management*, Vol. 36, No. 9 pp. 134–153.

Hillebrand, L., Raisch, S. and Schad, J. (2025) "Managing with Artificial Intelligence: An Integrative Framework", *Academy of Management Annals*, No. 19, pp. 343-375.

Karimi, J. and Walter, Z. (2016) "Corporate Entrepreneurship, Disruptive Business Model Innovation Adoption, and Its Performance: The Case of the Newspaper Industry", *Long Range Planning*, Vol. 46, No. 3, pp. 342-360.

Khan, M, Milisavljevic, S. J. and Salonitis, K. (2026), "Descriptive Maturity Model for Digital Servitization of SMEs", *Journal of Manufacturing Technology Management*, Vol. 37, No. 1 pp. 87–110.

Kohtamäki, M., Henneberg, S.C., Martinez, V., Kimita, K. and Gebauer, H. (2019) "A configurational approach to servitization: review and research directions", *Service Science*, Vol. 11, No. 3, pp. 213-240.

Kolagar, M., Parida, V. and Sjödin, D. (2022) "Ecosystem transformation for digital servitization: a systematic review, integrative framework, and future research agenda", *Journal of Business Research*, Vol. 146, pp. 176-200.

Kristiansen, M. and Aas, T.H. (2026) "Building legitimacy in digital servitization: an exploratory case study in manufacturing", *Journal of Manufacturing Technology Management*, Vol. 37 No. 1 pp. 65–86.

Laamanen, T., Weiser, A. K., von Krogh, G. and Ocasio, W. (2025) "Artificial intelligence in adaptive strategy creation and implementation: Toward enhanced attentional control in strategy processes", *Long Range Planning*, Vol. 58, No. 4, 102561.

Livieratos, A. D., W. Vanhaverbeke, G. Tsekouras, and Angelakis, A. (2024) "Unveiling the Black Box of Open Innovation in SMEs: Evidence-Based Key Success Factors and Challenges", *Journal of the Knowledge Economy*, Vol. 15, No. 2, pp. 1-24.

Marcon, E., Le Dain, M-A., Alejandro, G. and Frank, A. G. (2022) "Designing business models for Industry 4.0 technologies provision: Changes in business dimensions through digital transformation", *Technological Forecasting and Social Change*, Vol. 185, 122078.

Massa, S., Annosi, M. C., Marchegiani, L. and Messeni Petruzzelli, A (2023) "Digital Technologies and Knowledge Processes: New Emerging Strategies in International Business. A Systematic Literature Review", *Journal of Knowledge Managemen*t, Vol. 27, No. 11, pp. 330–387.

Müller, J. M., Buliga, O. and Voigt, K. I. (2021) "The role of absorptive capacity and innovation strategy in the design of industry 4.0 business Models-A comparison between SMEs and large enterprises", *European Management Journal*, Vol. 39, No. 3, pp. 333-343.

Ogink, R. H., Goossen, M. C., Romme, A. G. L. and Akkermans, H. (2023) "Mechanisms in open innovation: A review and synthesis of the literature", *Technovation*, Vol. 119, 102621.

Pham, C.M., Lokuge, S., Nguyen, T., Adamopoulos, A. (2024) "Exploring knowledge management enablers for blockchain-enabled food supply chain implementations", *Journal of Knowledge Management*, Vol. 28, No. 1 pp. 210-231.

Pauli, T., Marx, E., Fielt, E., Marheine, C. and Matzner, M. (2025) "Inhibitors and Enablers of Leverage in Digital Industrial Platform Ecosystems", *Business & Information Systems Engineering*, Vol. 67, No. 5, pp. 711-731.

Qin, L. and Li Sun, S. (2026) "Knowledge dependence in global value chains", *Research Policy*, Vol. 55, No. 2, 105392.

Spieth P., Schneckenberg D. and Matzler K. (2016) "Exploring the Linkage between Business Model & Innovation and the Strategy on the Firm", *R&D Management*, Vol. 46, No. 3, pp. 403-413.

Spithoven, A., Vanhaverbeke, W. and N. Roijakkers. (2013) "Open Innovation Practices in SMEs and Large Enterprises." Small Business Economics, Vol. 41, pp. 537–562.

Tate W.L., Ellram L.M. and Kirchoff F. (2010) "Corporate Social Responsability Reports: A thematic Analysis related to Supply Chain Management", *Journal of Supply Chain Management*, Vol. 46, No. 1, pp. 19-44.

Teece D.J. (2010) "Business Model, Business Strategy and Innovation", *Long Range Planning*, Vol. 43, Nn 2-3, pp. 172-194.

Teoh, M. F., Ahmad, N. H., Abdul-Halim, H. and Kan, W. H. (2023) "Digital business model innovation among small and Medium-sized enterprises (SMEs)", *Global Business and Organizational Excellence*, No. 42, pp. 5-18.

Thomas, L. D. W. and Ritala, P. (2022) "Ecosystem Legitimacy Emergence: A Collective Action View", *Journal of Management*, Vol. 48, No. 3, pp. 515-541.

The Cascade Paradox: Information Refraction in Hierarchical Organisations, AI-Driven Power Shifts, and the Future of Knowledge Cascades

Noémi Szilvia Lőrincz, Márta Kiss
Corvinus University of Budapest, Budapest, Hungary

Abstract

In knowledge-intensive organisations, the cascade - the sequential relay of strategic information through management layers - is the dominant mechanism by which corporate direction is translated into individual action. Yet cascade practices are rarely neutral conduits: information bends as it travels, shaped by the cognitive, structural, and political properties of each relay agent. This paper introduces information refraction as a theoretical construct to capture this systematic transformation and presents empirical evidence from two complementary datasets spanning the European automotive industry.

The first dataset comprises original mixed-method research within a large European automotive manufacturing organisation: 61 survey respondents stratified across four hierarchical levels (working level, first-line managers, middle managers, and senior managers / divisional heads), supplemented by three focus groups. The second dataset is a multi-organisational AI adoption survey with 37 respondents drawing respondents from OEMs, Tier 1 and Tier 2 suppliers, and adjacent industry actors across Europe, Japan, Germany, Hungary, Portugal, and the United Kingdom.

Together, these datasets reveal a cascade paradox: the actors most responsible for relaying strategy - middle managers - exhibit the lowest strategic comprehension and the highest self-reported limitations in cascade execution; while across the broader automotive ecosystem, AI adoption remains shallow (65% of organisations at pilot or early-exploration stage), fragmented (80%+ limiting AI to less than 20% of production processes), and impeded by identical barriers - lack of internal expertise (73% of respondents) and resistance to change (62%). We argue these two findings are structurally connected: the cascade problem and the AI readiness problem are expressions of the same underlying dynamic - organisations that cannot transmit knowledge faithfully cannot absorb transformative technologies effectively.

We advance a prospective argument with significant implications for knowledge management theory: AI-driven data centralisation will fundamentally alter the cascade's role as a power mechanism. For organisations that invest in AI literacy and knowledge infrastructure, the cascade paradox resolves through genuine democratisation - employees with strong analytical capacity gain direct data access, bypassing refraction. For those that do not, distortions compound. In an automotive sector navigating simultaneous disruption across electrification, supply chain fragility, and geopolitical volatility, cascade quality and AI readiness are not separate strategic priorities - they are the same priority.

Keywords: information cascade, knowledge management, organizational failure, middle management, upward communication

Paper type: Academic Research Paper

1 Introduction

Every organisation has a theory of how strategy becomes action. In most large corporations that theory is the cascade. An executive communicates direction to senior managers; senior managers translate it to middle managers; middle managers relay it to teams. The underlying assumption is elegant: hierarchical relay preserves fidelity while adding contextualisation at each layer.

The practice tells a different story: strategic initiatives lose clarity, critical information is delayed, simplified beyond utility, or withheld entirely. Employees act on incomplete signals. And when things go wrong, the diagnosis rarely identifies the cascade as the failure mode - because cascade dysfunction is invisible by design. It produces not loud errors but quiet erosions: of alignment, of trust, of competitive agility.

This paper calls this phenomenon the *cascade paradox*: the mechanism organisations depend upon most heavily to align human behaviour with strategic intent is simultaneously the mechanism most vulnerable to systematic distortion. We name the core dynamic information refraction - by analogy with optics, where light bends as it passes through media of different density. In organisations, information bends as it passes through layers of different authority, interest, cognitive load, and communicative competence.

The timing of this inquiry is not incidental. Across the automotive industry - among OEMs, Tier 1 suppliers, Tier 2 manufacturers, and adjacent actors - the pressure to adopt and leverage artificial intelligence is accelerating. Yet our multi-organisational survey reveals that AI adoption across European automotive actors remains predominantly shallow: most organisations are at pilot stage or early exploration, and the barriers they

cite - lack of internal expertise, resistance to change, poor data availability - are not primarily technological, they are organisational. They are, in a precise sense, cascade problems.

Our contribution is threefold. First, we present empirical evidence of cascade dysfunction across hierarchical levels within a large automotive manufacturing context, enriched by cross-industry AI readiness data spanning multiple OEMs and supplier tiers. Second, we theorise information refraction with greater conceptual precision than existing literature affords. Third, we argue that the rise of AI-enabled data access will not merely improve cascade efficiency - it will dismantle the cascade's role as the primary mechanism of strategic knowledge diffusion, with profound implications for where organisational power resides.

In the first theoretical section of this paper, we outline the key organizational contexts in which information generation, interpretation, and transmission play a critical role in decision-making. As organizations become increasingly complex and data-intensive, the volume and sophistication of available information grow rapidly, placing greater demands on hierarchical information-processing capacities. While advanced analytical tools and digital systems can enhance local decision-making, they also intensify coordination challenges across organizational levels. In such environments ensuring that relevant insights are accurately communicated upward under conditions of uncertainty and time pressure becomes a significant managerial challenge. Failures in this process can lead to delayed responses, strategic misalignment, and substantial performance losses. Through this

research, we seek to identify these mechanisms and assess how informational cascades shape organizational outcomes in complex operational settings. Furthermore, the paper devotes a whole chapter to important choice situations marked by high uncertainty and technological complexity, when informational cascades are most likely to occur. In such situations, middle management plays a critical role in interpreting, translating, and communicating specialist knowledge to top leadership. The chapter discusses how strategic narratives can eclipse localized technical discoveries, raising the possibility of distorted information flow. It also tackles the organizational challenges of managing risk, accountability, and uncertainty, such as responsibility, trust, and cross-hierarchical cooperation. Understanding these processes becomes increasingly crucial for avoiding strategic blind spots and organizational failure as technological and organizational complexity grows.

The major goal of this research is to discover the essential mechanisms that drive informational cascades within organizational hierarchies, with a focus on the role of middle management in shaping upward communication and strategic interpretation. Based on this purpose, the article addresses the following research question:

How do information cascades mediated by middle management influence upward information flow and strategic decision-making inside organizations?

This study uses a mixed-method research methodology to investigate how informational cascades influence organizational decision-making. It examines how middle management filters, transmits, and suppresses information across hierarchical levels using a combination of qualitative and quantitative methods. The paper aims to document patterns of information flow, upward communication, and decision-making across organizational hierarchies by distributing a questionnaire to the employees of selected major European firms. Additionally, it is possible to investigate the mechanisms, incentives, and limitations that lead to informational cascades and the repression of local knowledge within firms by conducting 3 focus groups with middle and senior managers. By clarifying how informational cascades arise and endure within hierarchical organizations, this work adds to the body of knowledge on organizational decision-making and information flow. The results provide useful information for companies looking to enhance strategic responsiveness and learning by emphasizing the influence of middle management on upward communication. Additionally, the study educates academics and managers on the dangers of systematic knowledge suppression and the circumstances that prevent local ideas from reaching high-level decision-makers. Lastly, it promotes more research into the consequences of informational cascades for long-term organizational success across organizational contexts, functions, and industries.

This work adds to the literature on organizational learning, decision-making, and information flow by expanding our understanding of informational cascades inside hierarchical organizations. It emphasizes how middle management influences the transfer and perception of knowledge during times of strategic uncertainty. The study also has practical implications for senior executives who want to create incentive structures and communication systems that lower the likelihood of knowledge suppression. Its drawbacks include a narrow focus on a small number of firms and managerial positions. Future study should use longitudinal designs and look at additional mediating factors as psychological safety, incentive alignment, and corporate culture.

2 Theoretical Background

Organizations seldom fail due to insufficient knowledge; they fail because pertinent information does not ascend promptly and accurately. Previous studies on informational cascades (Bikhchandani et al, 1992; Bikhchandani et al, 2024) indicate that individuals may withhold their private knowledge when social or hierarchical cues recommend an alternative course. Middle management plays a pivotal role inside businesses, serving as both interpreter and conduit of information between operational levels and senior leadership. Nonetheless, incentive frameworks, conformity constraints, and career risks frequently compel middle managers to comply to prevailing narratives instead of expressing opposing or adverse signals (Dvorak et al, 2025). Consequently, organizational learning failures and strategic blindness result from the systematic filtering, distortion, or silencing of local information (Haryjani et al, 2024). To understand how middle management leads to organizational failure, this paper looks at informational cascades.

Contemporary industries operate in environments characterized by rapid technological change, regulatory pressure, and increasing competitive intensity (Garrido-Moreno et al, 2024). Such conditions place significant strain on organizational information-processing and decision-making capabilities. As markets become more dynamic and complex, organizations face growing challenges in interpreting signals, coordinating responses, and aligning strategic actions across hierarchical levels. The influx of new entrants and disruptive innovations intensifies the need for timely and accurate upward communication. In this context, the ability of firms to adapt depends not only on access to information, but also on how effectively that information is filtered, interpreted, and escalated. Failures in these processes can result in strategic misalignment, delayed adaptation, and sustained competitive disadvantage (Ilmola-Sheppard & Kuusi, 2013). Not just the external environment has significant impact, but the increasing organizational complexity and data availability also have expanded the scope of managerial decision-making across operational and strategic domains. While advanced analytical tools can generate rich local insights, their organizational impact ultimately depends on how such information is interpreted, filtered, and communicated within hierarchical structures. A growing body of literature has examined information processing, sensemaking, and decision-making in organizations (Weick et al, 2005, Eckstein et al, 2025); however, comparatively little research offers analysis of how informational cascades emerge and persist at the middle management level. To address this gap, this paper reviews relevant literature on informational cascades and organizational decision-making and complements it with empirical evidence drawn from large organizations. By examining how middle managers translate and escalate information, the study sheds light on the prevalence and consequences of distorted information flow in complex organizational settings.

2.1 Information cascades in organisational hierarchies

The concept of information cascades emerged from economics and decision theory, describing situations in which individuals rationally discard private information in favour of publicly observable behaviour (Banerjee, 1992; Bikhchandani et al., 1992). Ribeiro et al. (2019) extended the framework to study message distortion in social diffusion chains, finding systematic content alteration as messages propagate without face-to-face interaction. Intra-organisational communication scholarship has addressed related phenomena under different labels: information filtering (Huber, 1991), knowledge stickiness (Szulanski, 1996), and communication distortion (O'Reilly & Roberts, 1974).

What unites these traditions is recognition that hierarchical position creates both motive and opportunity to alter information in transit. Huber and Daft (1987) documented that information richness declines with each relay step; Kiesler and Sproull (1982) observed systematic upward-filtered distortion in manager-to-executive communication. What is less studied - and what our research addresses - is the multi-level dynamics of downward strategic cascade, particularly in manufacturing organisations where cascade is formalised as an explicit management practice.

Nonaka and Takeuchi's (1995) knowledge creation model positions middle managers as 'knowledge engineers' who translate tacit executive vision into explicit operational knowledge. This framing treats the translation function as inherently enabling. Our data suggest it is also structurally vulnerable: translation capacity varies significantly across individuals and layers, and the conditions that support faithful translation - adequate information access, communicative competence, psychological safety to ask clarifying questions - are unevenly distributed.

2.2 The information refraction construct

We propose information refraction as a theoretical construct to describe the multi-dimensional transformation of information as it moves through organisational hierarchy. Distinct from distortion (which implies error) or filtering (which implies simple selection), refraction captures the systematic, structurally predictable bending of meaning induced by the properties of each hierarchical layer through which information passes. We define the *refractive index* of a managerial layer as a composite of four properties:

1. information access - the completeness and richness of what the layer receives;
2. translation capacity - the cognitive ability and communicative skill to convey strategic meaning faithfully;
3. political interest - what the layer gains or loses from transmitting specific content accurately; and
4. structural position - the layer's proximity to executive decision-making and exposure to multiple information streams. When any of these factors is suboptimal, refraction occurs: the message arrives at the next layer altered.

The refraction metaphor carries a further implication with direct relevance to AI disruption: if the medium through which information travels is removed - if employees gain direct, unmediated access to strategic data - the refractive effect disappears. The cascade collapses not because organisations abandon it deliberately, but because the technological conditions that made it necessary no longer obtain.

2.3 AI, data democratisation, and organisational power

A growing body of research examines how AI-mediated knowledge access reshapes organisational power dynamics. Gausen et al. (2023) developed a Consequence-Mechanism-Risk framework identifying that AI knowledge access systems interact non-trivially with existing power structures, potentially reinforcing asymmetries as much as dissolving them. Ide and Talamàs (2025) model the macroeconomic consequences of autonomous AI in knowledge hierarchies, finding that AI displaces humans towards complex problem-solving while restructuring the organisational demand for expertise at different levels. Recent theoretical work on GenAI and organisational structure argues that the technology's consequences for hierarchy depend critically on deployment design: the same technology can upskill or deskill depending on where in the hierarchy it is inserted. This resonates directly with our empirical finding that data access asymmetry - rather than strategic ambiguity - is the binding constraint on cascade fidelity in our research contexts. The technological intervention that dissolves the cascade paradox is not AI per se, but AI-enabled direct access to organisational intelligence.

3 Methodology

3.1 Mixed-method design

This study employs a concurrent mixed-methods design (Creswell & Clark, 2017), integrating quantitative survey evidence with qualitative focus group data, across two complementary research streams. The concurrent structure allows triangulation between within-organisation cascade dynamics (Research Stream 1) and industry-wide AI adoption patterns (Research Stream 2), enabling a theoretically integrated analysis that neither dataset alone could support.

Research Stream 1 was conducted within the European manufacturing division of a large global automotive OEM. The division encompasses several hundred members across multiple European locations. The organisation employs a formally institutionalised cascade communication practice, including regular all-hands briefings, daily team stand-up meetings, and structured one-to-one sessions at each hierarchical level. Survey data were collected across four stratified hierarchical groups, using anonymised digital instruments administered via Microsoft Forms. Table 1 presents the sample structure.

Table 1: Research Stream 1 - Survey Sample Structure

Hierarchical Level	n	Survey Instrument	Additional Qualitative Data
Working Level (< 5 yrs tenure)	7	Stream 1 Survey A	Focus Group 1
Working Level (≥ 5 yrs tenure)	13	Stream 1 Survey A	Focus Group 1
First-Line Managers	11	Stream 1 Survey B	Focus Group 2
Middle Managers	18	Stream 1 Survey B	Focus Group 2
Senior Managers / Divisional Heads	12	Stream 1 Survey C	Focus Group 3
TOTAL	61		3 Focus Groups

Three focus groups - corresponding to the working-level, managerial, and senior-managerial strata - explored lived cascade experience, friction points, and normative expectations of the cascade system. Key quantitative measures included: information reception adequacy (5-point Likert); strategic goal comprehension (5-point Likert); cascade confidence (categorical); and channel preference (multi-select).

Research Stream 2 was conducted to examine AI adoption in the automotive sector. The survey instrument covered AI technology portfolio, adoption maturity stage, departmental diffusion, key drivers and barriers, supply chain integration, production automation depth, quality management integration, staff training readiness, and sustainability/circular economy applications.
The sample (n = 37) was deliberately multi-organisational, spanning the automotive value chain rather than any single company. Respondents were drawn from OEMs (n = 26), Tier 1/2 car parts manufacturers (n = 7), and adjacent actors including professional services, maritime, and agri-manufacturing (n = 4). Geographically, respondents were headquartered across Belgium (n = 19), Japan (n = 5), Germany (n = 3), Hungary (n = 2), the UK (n = 1), Portugal (n = 1), Poland (n = 1), and Spain (n = 1). By role, respondents included management (n = 18), specialists (n = 9), C-suite (n = 7), and technical leads (n = 3). All organisations had 500+ employees with the exception of five mid-size organisations (101–500 employees).

4. Results

The most striking finding from Research Stream 1 is the systematic variation in strategic comprehension and information satisfaction across hierarchical levels. Table 2 presents mean scores across key dimensions.

Table 2: Stream 1 Survey Results by Hierarchical Level (n = 61)

Hierarchical Level	n	Info reception (mean /5)	Goal clarity (mean /5)	Can explain goals to next level
Working Level	20	3.80	3.85	Open text - see 4.2
First-Line Managers	11	3.73	4.27	45% Yes | 45% With limitations | 9% No
Middle Managers	18	3.89	3.78 ⚠	28% Yes | 61% With limitations | 11% No
Senior Managers / HO	12	4.67	4.58	75% Yes | 25% With limitations | 0% No

The data reveal a pronounced comprehension inversion at the middle manager level. Despite being the primary cascade relay point - the actors most responsible for translating strategy to working-level members - middle managers report both the lowest goal clarity (3.78/5) and the highest rates of self-reported limitation (61%) in their ability to explain goals to the next level. This is not a failure of individual capability but a structural phenomenon: middle managers receive partially-filtered information from senior layers, lack systematic access to the analytical context behind strategic decisions, and are expected to re-contextualise that information for operationally diverse teams - often without standardised tools or preparation time.

4.1 The supervisor as information gatekeeper

Working-level respondents identify the manager or supervisor as a primary information source in 65% of cases - the most cited single channel alongside the company portal. This dependence is not passive. Qualitative responses reveal active reliance on the supervisor as interpreter and contextualiser:

"Depending on the type of announcement, but in general I rely a lot on my supervisor for sharing company communication." (Working-level respondent, Vehicle Supply Chain)

This reliance creates a cascade vulnerability: when the gatekeeper's comprehension is limited - as our middle manager data shows - the information arriving at working level is doubly filtered: first by organisational distance from strategy, then by the gatekeeper's interpretation. Several respondents articulate an awareness of this structural inequity:

"If one supervisor is less sharing to employees and other supervisor more - it creates disadvantage in information available to some employees which is difficult to balance." (Working-level respondent, Strategy Planning)

This respondent has identified the core mechanism of information refraction empirically: systematic inequality in information availability as a function of the refractive properties of individual relay agents. Two employees at identical organisational levels in adjacent departments may operate with fundamentally different understandings of corporate strategy - not because the strategy is ambiguous, but because their respective relay agents bend it differently.

A further diagnostic emerges from Focus Group 1: the perception that strategic information sometimes arrives from external news sources before internal cascade. One respondent noted they were *"learning things via news about the company before it was even shared internally".* This is a classic cascade lag indicator - time-sensitive information circumventing the relay chain through external channels because internal relay is too slow.

4.2 The middle management paradox

Middle managers exhibit a distinctive pattern: their information reception score (3.89/5) is adequate - they receive information - but their goal clarity score (3.78/5) is the lowest in the sample, and their qualitative responses reveal why. They receive information stripped of the strategic context that would allow faithful re-contextualisation. As one middle manager respondent describes the situation: 'communications on different tasks - there is a lot of noise. Need to simplify messages, visualise objectives and link to the higher goals.' This is a precise description of refraction: the signal arrives amplified in volume but reduced in coherence, and the manager must perform signal extraction before re-transmission - a task for which they often lack adequate raw material.

Focus Group 2 (managerial level) surfaces the same dynamic from within: managers explicitly request standardised 'slide decks for complex topics' as cascade support tools, and advocate for 'consistency in messaging to avoid confusion.' These are not requests for more information - they are requests for better-structured information that reduces the refractive burden at the relay point.

One working-level respondent articulates the ideal from the receiving end: 'Be a filter but do not over-filter the information, assuming I don't need to know on my level.' This captures the fundamental tension of middle management's cascade role: filtering is necessary, but over-filtering is the precise mechanism by which strategic intent dies in transit.

4.3 Senior management: high fidelity, high distance

Senior managers and Heads of Division report the strongest outcomes: 4.67/5 for information reception, 4.58/5 for goal clarity, and 75% confident in explaining strategy to the next level. Structurally expected - they receive information closest to source. But the high-fidelity position carries a design assumption: that they will faithfully and completely transmit their superior comprehension downward. One senior respondent's comment challenges this assumption:

'Honestly, I don't know the divisional strategy. Was it shared with all members?' - a striking admission of cascade opacity even at senior level in one organisational unit.

Focus Group 3 (senior level) reveals a shared appetite for structural discipline: standardised cascade processes, supporting documents, and quality-controlled messaging. The irony is that the actors closest to strategic information are simultaneously the most articulate about the need for cascade infrastructure - infrastructure that the system currently lacks.

4.4 Integration: the cascade paradox in the age of AI

4.2.1 The Structural connection between cascade dysfunction and AI unreadiness

Read in isolation, our two datasets tell two different stories: one about how information bends in hierarchical relay chains, another about how AI adoption in automotive is slower and shallower than discourse suggests. Read together, they reveal a single structural dynamic.

Organisations that rely on cascade as their primary mechanism of strategic knowledge diffusion share three characteristics with organisations that struggle to absorb AI: they are information-asymmetric (different actors have access to different quality and volume of data); they are translation-dependent (meaning is not self-evident but requires human intermediation at each relay point); and they are politically mediated (what gets shared,

and how, reflects the interests of the relay agent as much as the intent of the sender). These three characteristics define both cascade dysfunction and AI adoption failure.

The barriers that 73% of our Stream 2 respondents cite - lack of internal expertise - is precisely what cascades were historically supposed to develop: translational capacity at each organisational layer. The cascade was the knowledge management solution. It is now also the knowledge management problem. And the same dynamic - inadequate translation at the point of relay - is what prevents AI outputs from becoming organisational intelligence.

4.2.2 The inflection point: what AI-enabled data access changes

We are at the threshold of a structural shift. The centralisation of organisational data in cloud platforms, combined with AI-enabled query and synthesis capabilities - large language models, real-time analytics, intelligent search, predictive dashboards - creates the technical precondition for a qualitatively different information architecture. When employees can query organisational data directly, contextualise it against their role, and generate actionable insight without managerial relay, the refractive medium is removed.

This shift creates what we term the cascade collapse scenario: not a catastrophic failure but a structural obsolescence of the cascade as the primary mechanism of strategic information diffusion. The cascade does not disappear - but it transforms from a necessity (the only way information can travel) into a choice (one channel among many, selected for its interpretive value rather than its positional monopoly).

In this emerging architecture, the power dynamics of organisational knowledge shift profoundly. Currently, internal politics partly function by controlling information flow: the actor who decides what gets shared, with whom, and in what frame holds structural power. In an AI-enabled environment with democratised data access, this power diminishes. Employees with high analytical capacity - regardless of hierarchical position - can navigate data directly, form evidence-based positions, and contribute strategic insight without requiring the relay chain's approval. The 'talent dividend' of AI accrues to those with cognitive and analytical depth, not those with privileged access to cascaded information.

4.5 Two scenarios: democratisation or amplified dysfunction

The theoretical implication splits into two divergent futures, contingent on organisational investment decisions being made now.

Scenario A - The democratisation dividend

Organisations that invest systematically in AI literacy, data infrastructure, and equitable access architecture will experience genuine knowledge democratisation. Working-level members with strong analytical capacity - currently constrained by the refractive chain above them - will be liberated to operate on the basis of direct data access. The cascade paradox resolves: the middle manager who currently struggles to explain divisional strategy because they lack analytical context will, in this scenario, have that context directly available. Their role shifts from information relay agent to interpretive partner - a qualitatively higher-value function.

For automotive OEMs and their supplier networks, this scenario represents a significant competitive acceleration. The speed at which supply chain disruptions are diagnosed, production decisions are made, and strategic pivots are executed depends on the density of intelligence throughout the organisation - not just at the top. Scenario A multiplies that density.

Scenario B - The amplified dysfunction

Organisations that fail to invest in AI readiness, data quality, and workforce capability will not experience the cascade problem dissolving - they will experience it amplifying. The arrival of AI tools that employees begin to use ad hoc, without governance or interpretive training, generates a new form of refraction: AI-mediated distortion. Data that is unstructured, ungoverned, and person-dependent - common in manufacturing environments with legacy systems - produces unreliable AI outputs. Employees who cannot evaluate those outputs critically propagate AI-generated misinformation through the same structural mechanisms as before, but faster and at greater scale.

Our Stream 2 data suggests that many organisations are currently positioned closer to Scenario B than Scenario A: 57% lack trained staff for AI output interpretation, QMS integration is absent in 52% of cases, and supplier AI collaboration is rejected by half the sample. These are not minor readiness gaps - they are the infrastructure preconditions for AI-enabled knowledge democratisation, and they are largely absent.

Table 3: Current vs. AI-Enabled Information Architecture - Structural Contrasts

Dimension	Cascade Architecture (Current)	AI-Enabled Architecture (Emerging)
Information flow direction	Sequential downward relay through layers	Simultaneous multi-directional; available at point of need
Interpretive authority	Held by relay agent (managers)	Distributed; AI provides synthesis layer
Speed of diffusion	Governed by relay chain frequency (weekly, monthly)	Near real-time to point of need
Quality determinant	Each relay agent's comprehension and fidelity	Data governance quality and AI literacy
Power mechanism	Information access asymmetry	Analytical and contextual capacity
Working-level position	Structurally dependent on supervisor interpretation	Can self-serve, self-contextualise
Middle management role	Information relay - translation under resource constraint	Interpretive partner - analytical collaborator
Primary failure mode	Refraction: meaning bent in transit	Amplification: unreliable AI outputs propagated at scale
Competitive implication	Alignment gaps slow strategic execution	AI literacy divide creates winner/loser split between orgs

4.6 Theoretical contributions

This paper makes three principal theoretical contributions. First, the information refraction construct provides a more analytically precise account of intra-organisational cascade dysfunction than prior notions of distortion or noise. By specifying the refractive index as a multi-dimensional property encompassing access, translation capacity, political interest, and structural position, we produce a framework that predicts where and why refraction occurs - not merely that it does. This operationalisability is the construct's key theoretical advantage.

Second, the middle manager paradox - the relay agent with the lowest strategic comprehension at the highest-volume cascade point - is a structural finding that challenges Nonaka and Takeuchi's (1995) enabling framing of middle management knowledge engineering. Middle managers are not simply translators; they are translators who receive systematically impoverished inputs and are expected to produce enriched outputs. The structural conditions for faithful translation are unevenly distributed, and the layer most responsible for the translation function is the layer most underserved by the system.

Third, the cascade collapse argument advances a prospective theoretical contribution: as AI-enabled data access becomes structurally available, the organisational theory of cascade becomes increasingly anachronistic. Knowledge management scholarship requires frameworks for the post-cascade organisation - in which the research question shifts from 'how do we relay information faithfully?' to 'how do we build organisational capacity to use information intelligently?' Our bi-dataset approach demonstrates how empirical cascade research and AI adoption research can be theoretically integrated rather than conducted in separate silos.

4.7 Practical implications for automotive organisations

For OEMs and their supplier networks, our findings generate three urgent areas of focus. First, organisations should diagnose their cascade architecture empirically - using stratified, cross-level surveys of the kind deployed here - before designing AI strategy. Cascade dysfunction is the condition into which AI will be deployed. Organisations that understand their cascade failure modes before AI adoption will design more resilient AI architectures. Those that do not will replicate their cascade problems at machine speed.

Second, cascade resilience interventions should target the middle manager layer as the highest-leverage point. This means providing direct access to the analytical material supporting strategic decisions (not just conclusions); standardising cascade support materials; and treating communicative and analytical competence as core management development priorities - not soft skills, but operational competencies.

Third, and most urgently, organisations must begin investing now in the AI literacy and data governance infrastructure that will determine which bifurcated scenario materialises. The window for proactive design is narrowing rapidly. Our data shows that the barriers to AI adoption are organisational, not technological - and organisational barriers are what leadership teams are uniquely positioned to address. Technical readiness without human readiness produces systems that cascade distortion at industrial scale.

5 Conclusions

The cascade is one of the most ubiquitous and least examined mechanisms in organisational life. Across a large European automotive manufacturing organisation and a multi-OEM, multi-supplier ecosystem spanning European automotive actors, our research reveals that cascade practice is simultaneously essential and systematically flawed - not because the concept is wrong, but because the structural conditions under which it operates reliably produce information refraction. Meaning bends as it travels through organisational layers, and the bending is not random: it is structured by the refractive properties of each relay agent, most critically at the middle management level.

The cascade paradox is both an empirical finding and a theoretical provocation. It demands that knowledge management scholars take seriously the political, cognitive, and structural dimensions of information relay within hierarchical organisations. It also demands temporal urgency: the cascade is not merely an inefficient relic of pre-digital management practice. It is the mechanism through which AI-generated intelligence will - or will not - reach the people and processes that can act on it.

Across the broader automotive ecosystem, our multi-organisational data reveals that AI adoption is shallow, human readiness is lagging, and the barriers are organisational rather than technological. This is not a pessimistic finding - it is a clarifying one. The organisations that will navigate the disruptions of electrification, supply chain fragility, and geopolitical volatility most successfully are not those that deploy the most sophisticated AI. They are the ones that build the human and structural capacity to use whatever AI they deploy - the capacity to receive information, understand it, act on it, and pass that understanding faithfully to others.

In the end, the cascade paradox and the AI readiness problem are the same problem: an organisation's knowledge is only as good as its weakest relay point. And the relay points that matter most are not the algorithms. They are the people.

6 Limitations and Future Research

While this paper represents a theoretically rich context, the cascade paradox's prevalence in other organisational types requires further investigation. The multi-organisational breadth compensates partially for this limitation but does not permit the depth of hierarchical analysis. The prospective AI argument is necessarily theoretical: the transformation it describes is in early stages of emergence. Our data establishes that most organisations are not yet positioned to realise the democratisation scenario. Longitudinal research tracking cascade dynamics as AI access is deployed at scale will be essential to test the bifurcated scenario predictions empirically.

7 Ethics Declaration

Prior to participation, all interviewees were provided with written information detailing the purpose of the study and the procedures for data handling. Informed consent was obtained from all participants based on the information provided. To ensure confidentiality, participants are represented anonymously and in a non-identifiable manner, in accordance with their stated preferences.

8 AI Declaration

In the development of this paper, the grammar correction and language refinement processes were supported by the use of DeepL Write. The tool assisted in enhancing the clarity, coherence, and fluency of the text without altering the substantive content or interpretation of the research findings. A critical review and editing process was undertaken by the authors to ensure the accuracy and alignment of the outputs with the intended meaning. We also used the Claud program for brainstorming and gathering reference materials. The final manuscript is a reflection of the authors' intellectual input and analytical interpretation.

References

Banerjee, A.V. (1992). A simple model of herd behavior. The Quarterly Journal of Economics, 107(3), 797–817.

Bikhchandani, S., Hirshleifer, D., & Welch, I. (1992). A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades. Journal of Political Economy, 100(5), 992–1026. http://www.jstor.org/stable/2138632

Bikhchandani, S., Hirshleifer, D., Tamuz, O. and Welch, I. (2024). Information Cascades and Social Learning. Journal of Economic Literature 62 (3): 1040–93. DOI: 10.1257/jel.20241472

Creswell, J.W. & Clark, V.L.P. (2017). Designing and Conducting Mixed Methods Research (3rd ed.). Sage Publications.

Dvorak, F., Fischbacher, U., Schmelz, K. (2025). Strategic Conformity or Anti-Conformity to Avoid Punishment and Attract Reward, The Economic Journal, Volume 135, Issue 666, February 2025, Pages 556–583, https://doi.org/10.1093/ej/ueae085

Eckstein, G., Shrestha, A. & Russo, F. (2025). Thirty-five years of sensemaking in the business & management research: a bibliometric analysis, review and discussion Manag Rev Q 75, 3191–3217 https://doi.org/10.1007/s11301-024-00458-5

Garrido-Moreno, A., Martín-Rojas, García-Morales, R. V. J. (2024). The key role of innovation and organizational resilience in improving business performance: A mixed-methods approach, International Journal of Information Management, Volume 77, 2024, 102777, ISSN 0268-4012,

Gausen, A., Mitra, B., & Lindley, S. (2023). A Framework for Exploring the Consequences of AI-Mediated Enterprise Knowledge Access and Identifying Risks to Workers. arXiv:2312.10076.

Hanafi, A.G., Hasim, S., Subri, N. I., Tahir, M. Z. (2024). The Impact of Middle Managers on Project Management Success: A Critical Analysis, Journal of Management Scholarship, 31 December 2024, https://doi.org/10.38198/JMS/3.2.2024.20

Hariyani, D., Hariyani, P., Mishra, S., Sharma, M. K. (2024). Causes of organizational failure: A literature review, Social Sciences & Humanities Open, Volume 10, 2024, 101153, ISSN 2590-2911, https://doi.org/10.1016/j.ssaho.2024.101153. IFKAD 2026 – Budapest, Hungary || 1-3 July 2026 Intelligent Knowledge For Sustainable Organizations

Huber, G.P. (1991). Organizational learning: The contributing processes and the literatures. Organization Science, 2(1), 88–115.

Huber, G.P. & Daft, R.L. (1987). The information environments of organizations. In F.M. Jablin et al. (Eds.), Handbook of Organizational Communication. Sage.

Ide, E. & Talamàs, E. (2025). Artificial Intelligence in the Knowledge Economy. arXiv:2312.05481.

Ilmola-Sheppard L, Kuusi O (2013). Information filters as one of the means of managing strategic fit in a complex environment. Foresight, Vol. 15 No. 2 pp. 132–151, doi: https://doi.org/10.1108/14636681311321130

Kiesler, S. & Sproull, L. (1982). Managerial response to changing environments: Perspectives on problem sensing from social cognition. Administrative Science Quarterly, 27(4), 548–570.

Nonaka, I. & Takeuchi, H. (1995). The Knowledge-Creating Company. Oxford University Press.

O'Reilly, C.A. & Roberts, K.H. (1974). Information filtration in organizations: Three experiments. Organizational Behavior and Human Performance, 11(2), 253–265.

Ribeiro, M.H., Gligorić, K., & West, R. (2019). Message distortion in information cascades. Proceedings of the World Wide Web Conference (WWW '19). ACM.

Szulanski, G. (1996). Exploring internal stickiness: Impediments to the transfer of best practice within the firm. Strategic Management Journal, 17(S2), 27–43.

Weick, K. E., Sutcliffe, K. M., & Obstfeld, D. (2005). Organizing and the Process of Sensemaking. Organization Science, 16(4), 409–421. https://doi.org/10.1287/orsc.1050.0133

Knowledge-Based Mechanisms of Sustainable Competitive Advantage under Environmental Pressure

Eduardo Sánchez-García[1], Javier Martínez-Falcó[1], Bartolomé Marco-Lajara[1], Matilda Shini[2], Carmen Einsle[1], Grazia Dicuonzo[2], Zhadyra Konurbayeva[3], Assel Sadenova[3]

[1] University of Alicante, Alicante, Spain

[2] University of Bari Aldo Moro, Bari, Italy

[3] D. Serikbayev East Kazakhstan Technical University, Kazakhstan

Abstract

Sustainability has emerged as a critical global challenge, compelling industries to integrate environmental considerations into their core strategies. The main objective of this study is to examine the direct influence of green transformational leadership and the mediating roles of green knowledge management and green motivation in achieving a sustainable competitive advantage. Primary data collected from 157 firms belonging to the mining sector are analyzed using PLS-SEM technique. The findings reveal that green transformational leadership significantly enhances sustainable competitive advantage, both directly and through the mediating mechanisms of green knowledge management and green motivation, which enable organizations to systematize environmental knowledge, aligning individual and organizational goals, and fostering proactive employee behaviors toward sustainability. These results contribute to clarify how leadership-driven sustainability visions are translated into organizational capabilities and extend the knowledge-based view by showing how sustainability contexts intensify firms' dependence on internal knowledge orchestration and motivational alignment. In industries where environmental challenges are unavoidable, competitive advantage is less about external positioning and more about the effective internal mobilization of knowledge resources. In this regard, this investigation contributes to sustainability research by shifting the focus from external compliance and outcomes to internal organizational processes that sustain competitive advantage over time. From a managerial perspective, the results suggest that firms seeking sustainable competitive advantage should invest in environmentally oriented leadership development, as well as in the design of robust green knowledge management systems and motivational practices, since leaders play a critical role in signalling the strategic importance of sustainability, but enduring advantages arise when environmental knowledge is institutionalized and employees are intrinsically motivated to apply it. For policymakers, the results underline the importance of supporting organizational capabilities related to leadership development and knowledge management, rather than relying exclusively on regulatory enforcement. This study shows that sustainable competitive advantage in environmentally sensitive industries is fundamentally a knowledge-based phenomenon, by integrating leadership, knowledge management, and motivation into a coherent framework, thereby offering new insights into how organizations can transform sustainability pressures into strategic opportunities through internal knowledge-driven mechanisms. It is concluded that leadership practices promoting environmental stewardship are essential in transforming operational frameworks to meet sustainability challenges and achieve a sustainable competitive advantage, leveraging organizational knowledge resources and fostering employee motivation to align the firms' environmental and economic objectives.

Keywords: Knowledge management, leadership, motivation, sustainability, PLS-SEM.

Paper type: Academic Research Paper.

1 Introduction

In the contemporary era, sustainability has emerged as a global challenge, capturing the attention of governments, businesses, and civil society alike, and the escalating impacts of climate change, resource depletion, and environmental degradation have underscored the urgent need for a shift toward sustainable development (Chomachaei *et al.*, 2024; Lyu *et al.*, 2024). Firms across industries are increasingly held accountable for their environmental footprints, facing regulatory pressures and heightened societal expectations to adopt practices that minimize harm to the natural environment, and the growing awareness of humanity's impact on ecosystems has prompted a fundamental reevaluation of traditional business practices, emphasizing the necessity of integrating environmental considerations into core organizational strategies, so that achieving entrepreneurial sustainability is essential for ensuring long-term economic viability and environmental stewardship (Ishaq *et al.*, 2024; Quttainah *et al.*, 2024). Positioned within the mineral-rich expanses of Kazakhstan, mining activity stands as one of the most important sectors of the nation's economy, blending its storied legacy with contemporary innovation to address the mounting global emphasis on

sustainability, making an important contribution to national GDP and global mineral markets, which are facing growing demands to integrate sustainable practices into its operations (Tazhibekova *et al.*, 2020; Atakhanova *et al.*, 2023; Kunarbekova *et al.*, 2024).

As one of the world's leading producers of essential minerals, Kazakhstan's mining sector plays a key role in driving the national economy as well as meeting global demands for resources, but these operations often come with significant environmental challenges, including resource depletion, habitat destruction, and greenhouse gas emissions which, coupled with increasing regulatory scrutiny and societal expectations for environmental responsibility, underscore the urgent need for transformative approaches to sustainability within the industry (Alimbaev *et al.*, 2020; Denissova *et al.*, 2023; Omurova *et al.*, 2024). Despite its significance, limited research has explored how leadership within the mining industry can catalyze sustainability outcomes and, specifically, the role of green transformational leadership, a leadership style that combines visionary guidance with a commitment to environmental principles, on fostering firms' sustainable competitive advantage remains underexplored (Sujati *et al.*, 2023; Zhong *et al.*, 2023; Hu *et al.*, 2024; Onifade *et al.*, 2024). This study seeks to address this gap by analyzing the role of green transformational leadership as a driver of sustainable competitive advantage in the mining sector, examining its direct influence and the mediating effects of green knowledge management and green motivation. Existing literature highlights the direct benefits of leadership on business sustainability but often fails to delve into the mechanisms that translate leadership into measurable sustainability outcomes (Razzaque *et al.*, 2024; Sajjad *et al.*, 2024).

2 Theory and Hypothesis

In this study, green knowledge management, which refers to the systematic sharing, integration, and application of sustainability-related knowledge within organizations, with the aim of enabling firms to innovate, optimize resource usage, and align their operations with environmental goals, and green motivation, which captures the intrinsic drive among employees to engage in environmentally responsible behaviors, are analyzed to provide insights into the main mechanisms through which sustainability-focused leadership can drive superior organizational outcomes (Ahmed *et al.*, 2021; Khan *et al.*, 2024). By fostering green knowledge management and green motivation, leaders can build resilient organizations capable of facing the dual imperatives of growth and sustainability, so that together, these variables can explain the ways in which green transformational leadership influences sustainable competitive advantage and builds long-term resilience within firms, by promoting shared values, and an organizational culture that prioritizes sustainability. Kazakhstan's mining industry, as a leading contributor to employment and foreign trade, provides an ideal context for exploring these dynamics, inasmuch as the sector's substantial environmental footprint necessitates a revaluation of its traditional business practices (Mahmood & Orazalin, 2017; Zarubin *et al.*, 2021; Raihan & Tuspekova, 2022).

Addressing these challenges requires leadership that fosters a culture of sustainability and, in this sense, the concept of green transformational leadership emerges as a transformative force that inspires organizations to adopt practices that harmonize economic and environmental goals, enhancing the capacity of enterprises to respond to regulatory pressures, innovate in resource efficiency, and reduce their environmental impact, acting as catalysts for embedding sustainability into the strategic and operational frameworks of mining firms (Tian *et al.*, 2023; Janjua *et al.*, 2024; Amegboleza & Ülkü, 2025) Furthermore, when leaders prioritize the dissemination of sustainability-related knowledge, they can create an environment where employees are equipped to understand the relevance of their actions to the organization's environmental goals, building the technical competencies necessary for implementing eco-friendly practices and fostering a deeper connection between employees' day-to-day tasks and the broader sustainability objectives of the firm, thus ensuring that sustainability is a tangible operational priority by establishing systems that effectively capture, share, and utilize environmental insights (Younis & Hussain, 2023; Shafait & Huang, 2024).

In connection with the above, this kind of leadership can influence the workforce by fostering a shared sense of purpose centered on environmental responsibility, so that through the reinforcement of shared values and the consistent integration of green principles, employees may perceive their roles as integral to the organization's success in achieving sustainability goals, understanding how their individual contributions connect to the broader environmental and organizational outcomes, which may affect their level of commitment, innovation, and proactive behaviors (Huang *et al.*, 2023; Ansari & Khan, 2024). The resource-based view underscores the significance of unique organizational resources, such as environmental knowledge and employee dedication, positioning environmental knowledge as a critical asset for improving operational efficiency and enabling strategic differentiation (Aslam *et al.*, 2024; Samadhiya & Agrawal, 2024). Building on this foundation, self-determination theory offers a complementary perspective by highlighting the role of motivation in fostering individual engagement and performance, unveiling the way in which leaders can inspire

employees to internalize organizational sustainability goals, cultivating a sense of ownership and personal responsibility for environmental initiatives, and creating an engaged workforce that actively seeks innovative solutions to environmental challenges, contributing meaningfully to the organization's sustainability agenda (Chong *et al.*, 2024; Patwary *et al.*, 2024).

Adding further depth to this framework, social identity theory provides valuable insights into the dynamics of collective action, establishing that individuals derive a sense of identity and belonging from their association with groups that embody shared values and objectives, so that, in the context of sustainability, employees who perceive their organization as genuinely committed to environmental stewardship may be more likely to develop a shared identity centered on these values, strengthening collective efforts, fostering collaboration, and enhancing the organization's ability to address complex environmental challenges through cohesive and unified action (Aggarwal & Agarwala, 2023; Rana & Arya, 2024). When applied to industries like mining, where the balance between operational efficiency and environmental sustainability presents unique challenges, these theoretical frameworks highlight the critical role of leadership-driven mechanisms to guide the integration of green knowledge management practices and fostering a shared green identity, facilitating the development of innovative solutions and strengthening a corporate culture that prioritizes sustainability as a strategic imperative, thus underscoring the transformative potential of leadership in shaping resilient organizations that can successfully face the complexities of a global economy increasingly defined by environmental consciousness.

Although previous studies have explored the influence of green transformational leadership on firms' sustainability, they have not sufficiently examined its influence on achieving a sustainable competitive advantage, nor how green knowledge management and green motivation mediate this relationship, particularly in environmentally sensitive sectors like mining (Özgül & Zehir, 2023; Xin & Wang, 2023; Abbas, 2024). This study seeks to fill this gap by providing empirical evidence from the mining industry in Kazakhstan, extending the understanding of how leadership, knowledge, and motivation interact to drive sustainability outcomes, offering a comprehensive view of how organizations can adapt to modern environmental challenges. Therefore, this article contributes to the literature by providing empirical evidence of the impact of green transformational leadership on sustainable competitive advantage, highlighting the mediating roles of green knowledge management and green motivation, challenging the existing paradigms by demonstrating the specific mechanisms through which leadership can drive sustainability outcomes in a context marked by significant environmental and economic stakes. For scholars, practitioners, and policymakers, these findings offer valuable insights into leveraging leadership and organizational dynamics to achieve superior environmental performance and a sustainable competitive advantage. Following this introduction, subsequent sections will present the theoretical background supporting the hypotheses, the methodology employed, the empirical analysis conducted, a discussion of the findings, and, finally, the conclusions, which synthesize the insights and offer actionable recommendations for practitioners and policymakers.

Based on the above, the following hypotheses are proposed:

- *Hypothesis 1 (+): Green transformational leadership has a positive effect on sustainable competitive advantage.*
- *Hypothesis 2 (+): There is a mediating effect of green motivation in the relationship between green transformational leadership and sustainable competitive advantage.*
- *Hypothesis 3 (+): There is a mediating effect of green knowledge management in the relationship between green transformational leadership and sustainable competitive advantage.*
- *Hypothesis 4 (+): Green knowledge management and green motivation have a double mediating effect in the relationship between green transformational leadership and sustainable competitive advantage.*

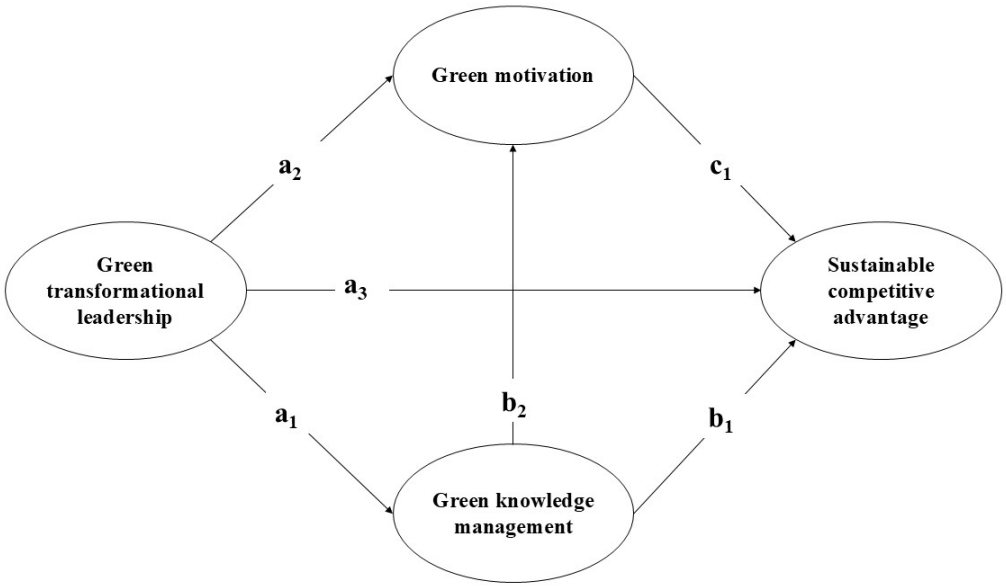


Figure 1. Nomogram of the proposed model

H1 = a3: Green transformational leadership → Sustainable competitive advantage; H2 = a2 x c1: Green transformational leadership → Green motivation → Sustainable competitive advantage; H3 = a1 x b1: Green transformational leadership → Green knowledge management → Sustainable competitive advantage; H4 = a1 x b2 x c1: Green transformational leadership → Green knowledge management → Green motivation → Sustainable competitive advantage.

3 Methodology

The mining industry in Kazakhstan is one of the most important for the national economy, contributing significantly to industrial and regional development while serving as a foundation for the growth of associated industries. According to the National Bureau of Statistics of the Republic of Kazakhstan, in 2023 the mining industry was composed of 1579 active firms, employing around 150,000 people, which is essential due to its high wages, which are more than double the national average, and the share of GDP from the mining industry in the industrial sector is 47%, and 35% in the total production of goods. Kazakhstan ranks among the top global producers of key mining resources, being the leading producer of uranium, with significant reserves of chromium, copper, gold, and zinc and, in terms of export, mining products accounted for nearly $10.5 billion in 2023, contributing 17.5% of Kazakhstan's total export value. Technological innovation is also shaping the sector, driven by large enterprises such as Kaz Minerals or Kazzinc, which have embraced Industry 4.0 principles, integrating AI for operational efficiency, predictive maintenance, and resource optimization, e.g., Kaz Minerals has achieved productivity rates ten times higher than similar operations globally by adopting automated systems and drone technology, and it has increased investments in copper, gold, and other metals aligned with global renewable energy trends suggest a promising trajectory (World Bank, 2023). This research focuses on the Kazakh mining sector as a context for exploring green transformational leadership and its effects on sustainable competitive advantage, with mediating roles of green knowledge management and motivation, providing interesting insights into sustainable strategies critical for enhancing the competitiveness and global positioning of the industry.

Data collection for this study involved distributing a questionnaire to Chief Executive Officers of mining firms in Kazakhstan during the period from September 9 to November 22, 2024, using the Qualtrics platform and Microsoft Outlook as dissemination tools. A total of 157 valid responses were collected, which corresponds to a response rate of 9.94%. The minimum sample size was determined using the minimum R^2 method as outlined by Hair *et al.* (2016). For a model with a maximum of three predictors and a minimum R^2 value of 0.480, a sample size of at least 38 was required. Additionally, calculations using the G*Power tool indicated that for significance levels of 5% and 1%, and a medium effect size, the minimum sample sizes needed were 74 and 108, respectively. The collected sample of 157 firms significantly exceeds these thresholds, enhancing the reliability of the results by providing a robust dataset that reduces the likelihood of statistical errors and increases the generalizability of the findings to the broader population of firms in the Kazakh mining sector.

To ensure consistency, reliability, and validity, the variables included in this study were measured using previously established scales. The independent variable, green transformational leadership, was assessed using six items from Chen and Chang (2013), while the dependent variable, sustainable competitive advantage, was evaluated with five items adapted from de Guimarães *et al.* (2018). The mediating variables, green knowledge management and green motivation, were measured using five and four items from Sahoo *et al.* (2023) and Úbeda-García *et al.* (2022). Data analysis was conducted using the PLS-SEM technique via the SmartPLS software, which is particularly suited for this study due to its ability to handle complex models with multiple constructs and indicators while providing robust predictive insights, being particularly valued for its accuracy and its focus on explaining variance in dependent variables (Hair *et al.*, 2012; Hair *et al.*, 2019a).

4 Results

The overall model fit was evaluated using the Standardized Root Mean Squared Residual (SRMR) yielding 0.072 < 0.08, aligning with the guidelines provided by Hu and Bentler (1998) and Hair *et al.* (2019a). Following this, the measurement models were assessed to ensure the constructs' reliability and validity, and to ensure the integrity of the research, discriminant validity is evaluated through Heterotrait-Monotrait ratio (HTMT). Moreover, using the methodology recommended by Hair *et al.* (2019b), the structural model was evaluated to provide a detailed and methodical examination of the theoretical constructs in the research framework, which involves analyzing direct, indirect, and total effects to confirm the proposed hypotheses and deepen the understanding of the relationships within the model. The model's predictive capabilities were assessed using R^2 values, with benchmarks of 0.25, 0.50, and 0.75 representing weak, moderate, and strong predictive power, respectively, as outlined by Hair *et al.* (2019a). The proposed model demonstrates robust explanatory power, accounting for 48.0%, 53.2%, and 63.5% of the variance in green knowledge management, green motivation, and sustainable competitive advantage, respectively, which validates the theoretical framework and emphasizes the critical interplay among environmental leadership, knowledge, and employee motivation to generate a sustainable competitive advantage.

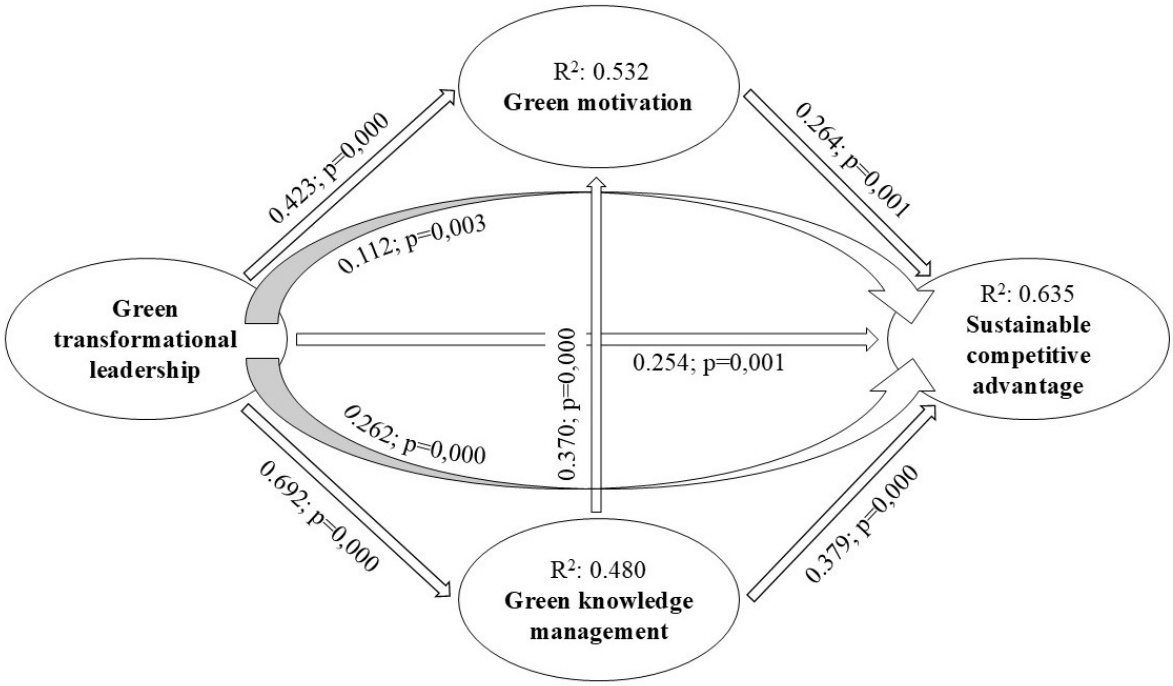


Figure 2. Direct and indirect relationships

This research offers significant contributions to the discourse on sustainable competitive advantage by integrating green transformational leadership within the framework of the Kazakh mining industry, increasing the understanding on how these firms can achieve a sustainable competitive advantage addressing their ecological footprint, and contributing to broader goals of sustainable development and economic resilience. The results evidence a positive and significant relationship between green transformational leadership and sustainable competitive advantage (H1), the mediating effects of green motivation (H2) and green knowledge management (H3), and the double mediating effect of these variables (H4). By exploring the role of green

transformational leadership alongside the mediating influences of green knowledge management and green motivation, it is established a way for integrating sustainability into core business strategies that transcends traditional operational models, which can offer firms a competitive edge while addressing environmental challenges in a global market increasingly defined by sustainability demands.

5 Conclusions

Sustainability has emerged as a key driver of long-term competitiveness and environmental responsibility in Kazakhstan's mining industry. This study emphasizes the pivotal role of green transformational leadership in embedding sustainability into the strategic and operational practices of mining companies, revealing that effective leadership fosters an environmentally conscious knowledge management system that aligns employees' motivation and efforts with broader sustainability goals, reshaping traditional business models to reduce environmental impact. The findings highlight how green knowledge management and green motivation mediate the link between leadership and sustainable competitive advantage, channeling the influence of leaders by systematically capturing, saving, and sharing environmental knowledge, and enhancing the motivation of the employees to act in an environmentally responsible manner, ensuring that their values align with organizational sustainability objectives, fostering a deeply engaged workforce committed to achieving green outcomes, which have a positive impact on the development of a sustainable competitive advantage, through their potential effects on innovation and operational efficiency.

Despite its valuable contributions, this study is not without its limitations. The reliance on cross-sectional data limits the ability to draw causal inferences about the relationships examined, so that future longitudinal studies could offer deeper insights into how these dynamics evolve over time. Moreover, the study focuses specifically on the mining industry in Kazakhstan, which, while offering a rich and relevant context, may limit the generalizability of the findings to other sectors or regions with different environmental and economic conditions. In addition, while the constructs of green transformational leadership, green knowledge management, and green motivation are explored comprehensively, other potentially relevant variables, such as organizational learning capacity or external stakeholder engagement, were not included, which could provide additional dimensions to the analysis. These limitations, however, do not diminish the study's contributions but rather highlight areas for future research to build upon.

6 Ethics declaration

This research is based on anonymized data and ethical approval is not required.

7 AI declaration

No AI tools were used to develop this research.

References

Abbas, J. (2024). Does the nexus of corporate social responsibility and green dynamic capabilities drive firms toward green technological innovation? The moderating role of green transformational leadership. *Technological Forecasting and Social Change*, *208*, 123698. https://doi.org/10.1016/j.techfore.2024.123698

Aggarwal, P., & Agarwala, T. (2023). Relationship of green human resource management with environmental performance: mediating effect of green organizational culture. *Benchmarking: An International Journal*, *30*(7), 2351-2376. https://doi.org/10.1108/BIJ-08-2021-0474

Ahmed, M., Guo, Q., Qureshi, M. A., Raza, S. A., Khan, K. A., & Salam, J. (2021). Do green HR practices enhance green motivation and proactive environmental management maturity in hotel industry?. *International Journal of Hospitality Management*, *94*, 102852. https://doi.org/10.1016/j.ijhm.2020.102852

Alimbaev, T., Omarova, B., Abzhapparova, B., Ilyassova, K., Yermagambetova, K., & Mazhitova, Z. (2020). Environment of East Kazakhstan: State and main directions of optimization. In *E3S Web of Conferences* (Vol. 175, p. 14008). EDP Sciences. https://doi.org/10.1051/e3sconf/202017514008

Amegboleza, A. A., & Ülkü, M. A. (2025). Advancing Sustainable Operational Efficiency in the Mining Industry: Trends, Innovations, Frameworks, and Future Research Directions. *Sustainable Development*, *1*, 1-22. https://doi.org/10.1002/sd.70211

Ansari, H. U. H., & Khan, S. N. (2024). Linking green transformational leadership and employee pro-environmental behavior: The role of intention and work environment. *Sustainable Futures*, *8*, 100336. https://doi.org/10.1016/j.sftr.2024.100336

Aslam, N., Shi, D., & Sahibzada, U. F. (2024). Combating green innovation in Italian luxury hotels: combination of social cognitive theory and natural resource-based view. *Business Process Management Journal*, *30*, 4, 1185-1212. https://doi.org/10.1108/BPMJ-07-2023-0518

Atakhanova, Z., & Azhibay, S. (2023). Assessing economic sustainability of mining in Kazakhstan. *Mineral Economics*, *36*(4), 719-731. https://doi.org/10.1007/s13563-023-00387-x

Chen, Y. S., & Chang, C. H. (2013). The determinants of green product development performance: Green dynamic capabilities, green transformational leadership, and green creativity. *Journal of business ethics*, *116*, 107-119. https://doi.org/10.1007/s10551-012-1452-x

Chomachaei, F., Gal-Or, E., Letizia, P., & Roma, P. (2024). The economic viability of the sharing economy business model and its environmental impact. *European Journal of Operational Research*, *315*(3), 1197-1209. https://doi.org/10.1016/j.ejor.2023.12.022

Chong, J. X., Gagné, M., Dunlop, P. D., & Wee, S. (2024). Facilitating newcomer motivation through internalization: A self-determination theory perspective on newcomer socialization. *Human Resource Management Review*, *34*(4), 101041. https://doi.org/10.1016/j.hrmr.2024.101041

de Guimarães, J. C. F., Severo, E. A., & de Vasconcelos, C. R. M. (2018). The influence of entrepreneurial, market, knowledge management orientations on cleaner production and the sustainable competitive advantage. *Journal of cleaner production*, *174*, 1653-1663. https://doi.org/10.1016/j.jclepro.2017.11.074

Denissova, O., Belgibayeva, A., & Suieubayeva, S. (2023). Evaluation of the Effectiveness of the "Green" Growth Policy Pursued at the Regional Level in the Republic of Kazakhstan. *Eurasian Journal of Economic and Business Studies*, *67*(4), 74-88. https://doi.org/10.47703/ejebs.v67i4.302

Hair, J.F.; Hult, G.T.M.; Ringle, C.M.; Sarstedt, M.; Castillo Apraiz, J.; Cepeda Carrión, G.; and Roldán, J.L. (2019a). *Manual de Partial Least Squares Structural Equation Modeling* (PLS-SEM), 1st ed. Madrid, ES: OmniaScience Scholar. https://doi.org/10.3926/oss.37

Hair, J.F., Babin, B.J., Anderson, R.E., & Black, W.C. (2019b). *Multivariate Data Analysis*, *8th ed.* Hampshire, UK: Cengage Learning.

Hair, J. F, Hult, G. T. M, Ringle, C, and Sarstedt, M. (2016). *A primer on partial least squares structural equation modeling* (PLS-SEM). California, USA: Sage Publications.

Hair, J. F., Sarstedt, M., Pieper, T. M., & Ringle, C. M. (2012). The use of partial least squares structural equation modeling in strategic management research: a review of past practices and recommendations for future applications. *Long range planning*, *45*(5-6), 320-340. https://doi.org/10.1016/j.lrp.2012.09.008

Hu, L. T., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological methods*, *3*(4), 424. https://psycnet.apa.org/doi/10.1037/1082-989X.3.4.424

Hu, K., Chen, Y., Mu, S., & Tan, Z. (2024). Green inclusive leadership, green policy, and pursuit of sustainability in natural resources: A firm-level analysis in context of Ukraine-Russia war. *Resources Policy*, *89*, 104607. https://doi.org/10.1016/j.resourpol.2023.104607

Huang, L., Guo, Z., Deng, B., & Wang, B. (2023). Unlocking the relationship between environmentally specific transformational leadership and employees' green behaviour: A cultural self-representation perspective. *Journal of Cleaner Production*, *382*, 134857. https://doi.org/10.1016/j.jclepro.2022.134857

Ishaq, M. I., Sarwar, H., Aftab, J., Franzoni, S., & Raza, A. (2024). Accomplishing sustainable performance through leaders' competencies, green entrepreneurial orientation, and innovation in an emerging economy: Moderating role of institutional support. *Business Strategy and the Environment*, *33*(2), 1515-1532. https://doi.org/10.1002/bse.3557

Janjua, N. A., Shi, D., & Sahibzada, U. F. (2024). Harnessing green innovation via green transformational leadership in Italian luxury hotels: key strategic takeaways. *International Journal of Hospitality Management*, *120*, 103739. https://doi.org/10.1016/j.ijhm.2024.103739

Khan, A. N., Mehmood, K., & Kwan, H. K. (2024). Green knowledge management: A key driver of green technology innovation and sustainable performance in the construction organizations. *Journal of Innovation & Knowledge*, *9*(1), 100455. https://doi.org/10.1016/j.jik.2023.100455

Kunarbekova, M., Yeszhan, Y., Zharylkan, S., Alipuly, M., Zhantikeyev, U., Beisebayeva, A., ... & Azat, S. (2024). The state of the art of the mining and metallurgical industry in Kazakhstan and future perspectives: A systematic review. *ES Materials & Manufacturing*, *25*, 1219. http://dx.doi.org/10.30919/esmm1219

Lyu, Y., Bai, Y., & Zhang, J. (2024). Green finance policy and enterprise green development: Evidence from China. *Corporate Social Responsibility and Environmental Management*, *31*(1), 414-432. https://doi.org/10.1002/csr.2577

Mahmood, M., & Orazalin, N. (2017). Green governance and sustainability reporting in Kazakhstan's oil, gas, and mining sector: Evidence from a former USSR emerging economy. *Journal of cleaner Production*, *164*, 389-397. https://doi.org/10.1016/j.jclepro.2017.06.203

Omurova, J., Zhanbossinova, A., Kurumbaeva, G., Kurbanova, N., Khassenova, Z., Bekboeva, A., & Karasartova, C. (2024). On the state of the environment and natural environment of the East Kazakhstan region. In *E3S Web of Conferences* (Vol. 539, p. 01004). EDP Sciences. https://doi.org/10.1051/e3sconf/202453901004

Onifade, M., Zvarivadza, T., Adebisi, J. A., Said, K. O., Dayo-Olupona, O., Lawal, A. I., & Khandelwal, M. (2024). Advancing toward sustainability: The emergence of green mining technologies and practices. *Green and Smart Mining Engineering*, *1*(2), 157-174. https://doi.org/10.1016/j.gsme.2024.05.005

Özgül, B., & Zehir, C. (2023). Top management's green transformational leadership and competitive advantage: the mediating role of green organizational learning capability. *Journal of Business & Industrial Marketing*, *38*(10), 2047-2060. https://doi.org/10.1108/JBIM-01-2022-0043

Patwary, A. K., Rasoolimanesh, S. M., Hanafiah, M. H., Aziz, R. C., Mohamed, A. E., Ashraf, M. U., & Azam, N. R. A. N. (2024). Empowering pro-environmental potential among hotel employees: insights from self-determination theory. *Journal of Hospitality and Tourism Insights*, *7*(2), 1070-1090. https://doi.org/10.1108/JHTI-10-2023-0713

Quttainah, M. A., & Ayadi, I. (2024). The impact of digital integration on corporate sustainability: Emissions reduction, environmental innovation, and resource efficiency in the European. *Journal of Innovation & Knowledge*, *9*(3), 100525. https://doi.org/10.1016/j.jik.2024.100525

Raihan, A., & Tuspekova, A. (2022). Role of economic growth, renewable energy, and technological innovation to achieve environmental sustainability in Kazakhstan. *Current Research in Environmental Sustainability*, *4*, 100165. https://doi.org/10.1016/j.crsust.2022.100165

Rana, G., & Arya, V. (2024). Green human resource management and environmental performance: mediating role of green innovation–a study from an emerging country. *foresight*, *26*(1), 35-58. https://doi.org/10.1108/FS-04-2021-0094

Razzaque, A., Lee, I., & Mangalaraj, G. (2024). The effect of entrepreneurial leadership traits on corporate sustainable development and firm performance: a resource-based view. *European Business Review*, *36*(2), 177-200. https://doi.org/10.1108/EBR-03-2023-0076

Sahoo, S., Kumar, A., & Upadhyay, A. (2023). How do green knowledge management and green technology innovation impact corporate environmental performance? Understanding the role of green knowledge acquisition. *Business Strategy and the Environment*, *32*(1), 551-569. https://doi.org/10.1002/bse.3160

Sajjad, A., Eweje, G., & Raziq, M. M. (2024). Sustainability leadership: An integrative review and conceptual synthesis. *Business Strategy and the Environment*, *33*(4), 2849-2867. https://doi.org/10.1002/bse.3631

Samadhiya, A., & Agrawal, R. (2024). Total productive maintenance and sustainability performance: resource-based view perspective. *Benchmarking: An International Journal*, *31*(7), 2177-2196. https://doi.org/10.1108/BIJ-10-2022-0635

Shafait, Z., & Huang, J. (2024). Examining the impact of sustainable leadership on green knowledge sharing and green learning: Understanding the roles of green innovation and green organisational performance. *Journal of Cleaner Production*, *457*, 142402. https://doi.org/10.1016/j.jclepro.2024.142402

Sujati, W., Yudoko, G., & Okdinawati, L. (2023). Transactional and transformational leadership styles and their impact on employees' acceptance of predictive maintenance analytics: Evidence from an Indonesian mining company. *IEEE Access*, *11*, 49675-49688. https://doi.org/10.1109/ACCESS.2023.3277006

Tazhibekova, K., Shametova, A., Urazbekov, A., Akhmetzhanov, B., Akenov, S., & Tulupova, S. (2020). Enhancing eco-economic efficiency of mineral deposit exploration to achieve sustainable development in the mining industry of Kazakhstan. *Progress in Industrial Ecology, an International Journal*, *14*(3-4), 212-228. https://doi.org/10.1504/PIE.2020.113425

Tian, H., Siddik, A. B., Pertheban, T. R., & Rahman, M. N. (2023). Does fintech innovation and green transformational leadership improve green innovation and corporate environmental performance? A hybrid SEM–ANN approach. *Journal of Innovation & Knowledge*, *8*(3), 100396. https://doi.org/10.1016/j.jik.2023.100396

Úbeda-García, M., Marco-Lajara, B., Zaragoza-Sáez, P. C., Manresa-Marhuenda, E., & Poveda-Pareja, E. (2022). Green ambidexterity and environmental performance: The role of green human resources. *Corporate Social Responsibility and Environmental Management*, *29*(1), 32-45. https://doi.org/10.1002/csr.2171

World Bank (2023). *Kazakhstan: Strategic Review of the Mining and Metallurgy Sector*. Retrieved on November 15/2024 from https://documents.worldbank.org/en/publication/documents-reports/documentdetail/402311468775805892/kazakhstan-strategic-review-of-the-mining-and-metallurgy-sector

Xin, C., & Wang, Y. (2023). Green intellectual capital and green competitive advantage in hotels: The role of environmental product innovation and green transformational leadership. *Journal of Hospitality and Tourism Management*, *57*, 148-157. https://doi.org/10.1016/j.jhtm.2023.10.001

Younis, Z., & Hussain, S. (2023). Green transformational leadership: bridging the gap between green HRM practices and environmental performance through green psychological climate. *Sustainable Futures*, *6*, 100140. https://doi.org/10.1016/j.sftr.2023.100140

Zarubin, M., Statsenko, L., Spiridonov, P., Zarubina, V., Melkoumian, N., & Salykova, O. (2021). A GIS software module for environmental impact assessment of the open pit mining projects for small mining operators in Kazakhstan. *Sustainability*, *13*(12), 6971. https://doi.org/10.3390/su13126971

Zhong, J., Shao, X., Xiao, H., Yang, R., & An, X. (2023). The research on the green leadership: a systematic review and theoretical framework. *Environment, Development and Sustainability*, *1*, 1-32. https://doi.org/10.1007/s10668-023-03960-0

Integrating Farmers' Behaviors into Agricultural Index-Based Insurance: A Knowledge-Based Theoretical Framework

Giulio Paolo Agnusdei[1], Donato Longo[2,3], Zahra Ardakani[2], Pier Paolo Miglietta[2,3]
[1]Pegaso University, Naples, Italy
[2]University of Salento, Lecce, Italy
[3]National Centre for HPC, Big Data and Quantum Computing, Bologna, Italy

Abstract

Agricultural index-based insurance has emerged as an important mechanism for managing climate-related risks; however, existing frameworks largely treat farmers' behavior as exogenous, overlooking how knowledge-driven adaptive responses to early-warning signals may dynamically mediate crop losses. This study proposes behavior-based index insurance (BBII), a theoretical framework that extends multi-hazard index-based insurance models by incorporating a behavioral compliance component into actuarial pricing. Specifically, the framework introduces an adaptive behavior index (ABI) as a composite metric that captures farmers' behaviors in response to early-warning signals across three dimensions: compliance with recommended actions, response timing, and intervention intensity. By integrating the ABI into the vulnerability module, BBII allows yield sensitivity to vary as a function of farmers' behavior, rather than relying on the multi-hazard index alone, thereby generating a dynamic tick value that scales payouts per unit of multi-hazard index variation according to behavior-adjusted yield sensitivity. This mechanism results in a behavior-sensitive premium structure in which premiums decrease with higher adaptive behavior and increase under low compliance. The proposed framework reconceptualizes insurance as an active risk-governance mechanism rather than a passive compensation mechanism, creating financial incentives for proactive adaptive behavior. By endogenizing adaptive behavior within index insurance design, this framework advances a dynamic, knowledge-driven approach to agricultural risk management that aligns pricing incentives with effective climate adaptation.

Keywords: index-based insurance; adaptive behavior index; climate risk management; early-warning systems; actuarial pricing.

Paper type: Academic Conceptual Paper

1 Introduction

Climate change poses significant and escalating challenges to agriculture, threatening food security, rural livelihoods, and national economies (Raza et al., 2019; Janni et al., 2024; Saleem et al., 2025). Temperature and precipitation, the fundamental climatic variables, regulate the entire crop growth cycle, from planting to harvest (Agnolucci et al., 2020). Deviations beyond optimal thresholds of these variables within each specific crop growth cycle can induce physiological stress, disrupt plant development, and ultimately reduce crop yields (Aslam et al., 2021; Moore et al., 2021; Chen et al., 2022; Angon et al., 2024; Muffler et al., 2024; Vandana et al., 2024; Zhang et al., 2025). Beyond direct physiological impacts, climate change may also alter planting and harvesting schedules, introducing greater uncertainty into farm-level decision-making and undermining the predictability upon which agricultural systems depend (Shah et al., 2024; Battaglia et al., 2025; Saleem et al., 2025).

As climate variability intensifies, adopting structured risk management strategies and decision support systems has become an urgent priority, enabling farmers to make appropriate decisions under current and future climate conditions (Menalled et al., 2026). Agricultural climate risk management is a multidisciplinary process encompassing the identification, assessment, and response to climate-related risks, with the ultimate aim of enhancing resilience and reducing potential damage to farming systems and rural livelihoods (Selvaraju et al., 2012). At the farm level, risk reduction measures, such as soil and water management, operationalize adaptation by embedding climate responsiveness directly into practice (Saleem et al., 2025). At the institutional level, effective risk management requires broad stakeholder collaboration, engaging agricultural extension services, meteorological agencies, financial institutions, and local cooperatives, to ensure that climate information reaches farmers in a real-time, credible, and actionable form (Booth et al., 2020; Parnell and Gangwish, 2023).

Agricultural insurance is a critical risk management mechanism to support and stabilize farmers' income (Chen et al., 2025; Hou and Wang, 2025). Despite its potential, traditional indemnity-based insurance models remain poorly suited to agricultural systems due to the systemic nature of climate-related risks, difficulties in verifying individual losses, and high transaction and claim-assessment costs (Sgroi and Sciancalepore, 2022;

Mateos-Ronco, 2025), limitations that are becoming increasingly evident (Ardakani et al., 2026). In response, index-based insurance models have emerged as a more tractable financial instrument, circumventing these limitations by linking payouts to objectively measurable climate indices rather than individually verified losses (Gesualdo et al., 2024; Zhu et al., 2025). Accordingly, a recently developed multi-hazard index-based insurance model represents a structured advancement, connecting climate-related risks to financial risk through hazard indices, vulnerability estimation, and actuarial pricing mechanisms (Ardakani et al., 2026).

While the multi-hazard index-based insurance model offers a robust representation of hazard intensity, severity, and exposure, it relies on a largely exogenous and static treatment of vulnerability, thereby overlooking the role of farmers' behavior and adaptive responses to early-warning signals. This represents a consequential gap, as farmers' decisions regarding input management, irrigation timing, and harvesting schedules directly mediate the relationship between climate-related risks and realized losses yet are not reflected in existing index construction or pricing frameworks. Recent advances in sensor-based agricultural decision systems demonstrate the feasibility of monitoring field and crop conditions in near real time at decision-relevant scales (Capone et al., 2025). In parallel, behavior-based pricing models, widely applied in usage-based insurance, provide a theoretical basis for incorporating behavioral compliance into premium calculations (Tselentis et al., 2016; Li et al., 2025). Together, these developments indicate the emergence of a new generation of index-based insurance models in which risk, vulnerability, and pricing are jointly shaped by hazard exposure and adaptive behavior.

Therefore, this study aims to theoretically extend the multi-hazard index-based insurance model (Ardakani et al., 2026) by incorporating a behavioral compliance component that accounts for farmer responses to climate-related early-warning signals. Specifically, the proposed framework introduces a behavior-based adjustment into the actuarial pricing structure, whereby premium calculations reflect not only hazard exposure, but also the degree to which farmers adopt recommended adaptive practices following the issuance of alerts. Hence, the model reframes insurance not merely as a passive compensation mechanism, but as an active risk-governance tool that creates financial incentives for proactive adaptation and thereby strengthens the link between early-warning systems and on-farm decision-making. This theoretical contribution is particularly relevant in the context of advancing sensor technologies and real-time farm monitoring, which increasingly make behavioral compliance observable and verifiable at scale. The proposed framework thus carries implications not only for actuarial design, but also for agricultural policy, extension services, and the broader institutional architecture of climate risk management.

The remainder of the paper is organized as follows. Section 2 reviews the role of data-driven decision-support systems and smart agricultural technologies in agricultural risk management, with emphasis on data availability, advanced sensing and analytics. Section 3 introduces the proposed theoretical framework that incorporates farmer behavior into index-based insurance design, presents its mathematical formulation, and discusses implications for actuarial pricing. Section 4 concludes by outlining implications for insurance design and policy, and by identifying directions for future empirical validation and practical implementation.

2 Background: Smart Technologies for Agricultural Risk Management

Resilient agriculture requires cooperative and coordinated actions across numerous sectors, while farmers' behaviors and decisions remain a particularly critical determinant (Savari et al., 2023; Song et al., 2025). Farmers' timely and appropriate responses to emerging climate signals can substantially reduce crop losses (Minoli et al., 2022; Ricart et al., 2023). Translating climate signals into on-farm actions and practices, however, is not straightforward; it depends on farmers having access to real-time, relevant, and actionable guidance, which underscores the need for effective decision support systems (Palutikof et al., 2019; Constable et al., 2022). A decision support system is an interactive computer-based platform that assists decision-making by integrating data, models, and tools (Rinaldi and He, 2014). In agriculture, decision support systems assist farmers in selecting appropriate practices and strategies to improve productivity, efficiency, and sustainability (Zhai et al., 2020; Fang et al., 2021; Mora et al., 2018; Di Gangi et al., 2023; Behera et al., 2025; Goud et al., 2025). The effectiveness of agricultural decision support systems therefore depends fundamentally on the availability of high-quality and timely data (Fang et al., 2021).

Data are the core operational input for decision support systems and a fundamental requirement for data-driven decision-making (Power, 2008). In agriculture, the availability of timely data is particularly critical due to the sector's strong dependence on environmental and climatic conditions, under which climate variability directly affects crop growth, yield, and resource efficiency (Mishra & Mishra, 2024). As climate variability intensifies, traditional intuition-based agricultural practices are increasingly inadequate for managing production risks and uncertainty (Arora, 2019; Kumar et al., 2023). In response, the agricultural sector is progressively shifting toward data-driven practices that rely on systematic data collection and analytical tools to

enhance productivity, efficiency, and resilience (MacPherson et al., 2022; Araújo et al., 2023). However, the true value of agricultural data lies in its ability to provide actionable insights that support timely, informed, and resilient decision-making (Carletto, 2021; Leal et al., 2024).

Smart agriculture is increasingly relying on advanced data collection and processing methods to enhance decision-making and operational efficiency (Basnet and Bang, 2018; Gamage et al., 2024; Soussi et al., 2024). Emerging data-collection technologies such as the Internet of Things (IoT), unmanned aerial vehicles (UAVs), satellite-based remote sensing, and robotic sensing platforms are enabling high-resolution, real-time monitoring of crop, soil, and environmental conditions (Padhiary et al., 2025). IoT sensor networks consist of distributed, connected sensors deployed in the field to continuously collect soil, crop, and environmental data in real time (Mansoor et al., 2025). UAVs are aerial platforms equipped with imaging sensors that provide high-resolution, flexible, and timely observations of crop health, soil conditions, and pest infestations (Xing et al., 2025). Satellite-based remote sensing (earth observation) uses satellites to deliver large-scale, repetitive measurements for monitoring crop growth, land use, and environmental variability (Nakalembe et al., 2021). Robotic sensing platforms are autonomous ground-based systems that integrate proximal sensors and vision technologies for detailed, in-field crop assessment (Łągiewska et al., 2025).

Data analysis and analytics techniques are subsequently essential for transforming raw data into actionable decisions (Sarker, 2021). Data analysis focuses on examining, cleaning, transforming, and interpreting data to understand historical or current phenomena by identifying patterns, trends, and relationships. It is typically descriptive or diagnostic and aims to explain what has happened or why it occurred (Devore, 2007; Gudivada, 2025). In contrast, data analytics is a broader and more comprehensive discipline that includes data analysis as a core component while extending beyond it toward prediction and decision support (Basnet and Bang, 2018; Gudivada, 2025). Rather than solely interpreting past data, data analytics emphasizes forecasting future outcomes and recommending actions, thereby supporting proactive and data-driven decision-making across complex systems (Abdul-Azeez et al., 2024). Moreover, data analytics techniques integrate advanced computational methods, automation, and technologies such as machine learning (Rejeb et al., 2025) and artificial intelligence (Mishra and Mishra, 2024) to analyze large and complex datasets on a continuous basis (Waqas et al., 2025).

Artificial Intelligence of Things (AIoT) integrates advanced sensing infrastructures with data analytics and artificial intelligence to enable learning, prediction, and adaptive decision-making (Benfradj et al., 2024). Unlike traditional IoT systems that primarily support data collection, AIoT architectures can interpret high-frequency sensor data and generate context-aware responses through machine-learning-driven analytics (Singh & Krishnamurthi, 2024). In agriculture, AIoT underpins precision farming by combining real-time monitoring of soil, crops, and environmental conditions with advanced analytics to support timely management decisions (Chowdhury et al., 2025). From a risk-management perspective, these capabilities strengthen early-warning systems and enable the observation of adaptive farmer responses, thereby providing a practical foundation for frameworks that move beyond static vulnerability representations and motivate the integration of behavior into index-based insurance design.

3 Methodology: Behavior-Based Index Insurance (BBII)

Behavior-based index insurance (BBII) extends the multi-hazard index-based insurance model (Ardakani et al., 2026) by incorporating farmers' behavior into the actuarial pricing mechanism. In contrast to the multi-hazard index-based insurance model, where payouts depend exclusively on a climate-related index, BBII integrates adaptive behavior into the vulnerability component of the insurance model. This allows insurance pricing to reflect both hazard exposure and farmers' responses to early-warning alerts. The BBII framework consists of four integrated modules: a hazard module, a behavioral module, a modified vulnerability module, and a financial pricing module. The hazard component remains unchanged from the original index construction (Ardakani et al., 2026), while behavioral adaptation is introduced through an adaptive behavior index (ABI), which interacts with hazard intensity to modify yield sensitivity and risk exposure.

3.1 Hazard module

The hazard module defines the multi-hazard index (Equation 1), which aggregates the effects of multiple climate hazards across different crop growth stages. The index captures the deviation of observed hazard levels from their respective critical thresholds and assigns stage-specific importance weights reflecting crop sensitivity during phenological development (see Ardakani et al., 2026 for details of the index construction process).

$$MHI = \sum_{i}^{n}\sum_{s}^{m} w_{is} \times (H_{is} - H_{critical}) \quad (1)$$

where *MHI* denotes the multi-hazard index, w_{is} and H_{is} represent the stage-specific weight and the value of climate hazard *i* during growth stage *s*, respectively. $H_{critical}$ refers to the critical threshold associated with each hazard.

3.2 Behavioral module

To explicitly account for behavioral response dynamics within the proposed insurance framework, this study introduces a behavior-dependent index, which is designed to quantify farmers' real-time responses to early-warning alerts generated by digital decision-support systems. While composite indices have been widely used to summarize farmers' adaptive behavior by aggregating multiple actions into a single index, existing applications largely focus on ex-post adoption of adaptation measures rather than on process-oriented responses to warnings (Song and Zhu, 2025). Early-warning systems are fundamentally action-oriented: their effectiveness depends not only on the accuracy of hazard detection, but also on how end users process warnings and translate them into timely and appropriate actions (Coffey et al., 2015). Recent warning-response research conceptualizes behavior as a multi-dimensional construct in which information reception, response timing, and intensity of action jointly determine risk-reduction outcomes (Xu et al., 2026).

Accordingly, to incorporate adaptive decision-making into the insurance framework, a behavior-dependent index, named the ABI, is introduced. The ABI quantifies the degree to which farmers respond to early-warning alerts issued by digital decision-support systems. Behavioral performance is measured based on three complementary dimensions: compliance with recommended actions following an alert, representing whether farmers implement the prescribed intervention, response timing, capturing the promptness of action relative to the issuance of the warning, and the intervention adequacy, reflecting whether the scale or intensity of the response is sufficient given the anticipated severity of the hazard. Mathematically, the ABI is defined as follows (Equation 2):

$$ABI_t = \sum_{i}^{n} w_i(R_{it} \times T_{it} \times V_{it}) \quad (2)$$

where *ABI* represents the aggregated behavioral score at time *t*, R_{it} denotes the degree of compliance with recommended actions for hazard *i*, T_{it} measures the timeliness of the response relative to the warning issuance, and V_{it} reflects the adequacy of intervention intensity. The parameter w_i denotes a hazard-specific weight. The multiplicative structure of Equation 2 reflects the conceptual assumption that effective adaptive response requires the simultaneous satisfaction of all three behavioral conditions: compliance without timeliness, or prompt action without adequate intensity, is assumed to result in diminished protective effectiveness. The ABI values close to one indicate highly responsive and timely adaptation behavior.

3.3 Vulnerability module

The vulnerability module describes the relationship between crop yield, climate hazards, and farmers' responses. Unlike previous formulations (Ardakani et al., 2026) that assume a fixed yield response to the climate hazard index, the BBII framework allows the sensitivity of yield to farmers' behavior rather than the hazard exposure alone. Accordingly, yield is modeled as (Equation 3):

$$Y_t = \beta_0 + \beta_1 MHI_t + \beta_2 ABI_t + \beta_3 (MHI_t \times ABI_t) + \epsilon_t \quad (3)$$

where *Y* denotes observed crop yield at time *t*, MHI_t represents the multi-hazard index (i.e., cumulative hazard intensity and severity), and ABI_t captures farmers' behavioral adaptation. The interaction term reflects the moderating effect of farmers' behavior on yield sensitivity to hazard exposure. The parameters β_1, β_2, and β_3 represent the marginal effects of hazard intensity, behavioral adaptation, and their interaction, respectively, while ϵ_t represents the error term.

The inclusion of the interaction term allows the marginal impact of climate hazard index on yield to vary as a function of farmers' behavioral adaptation. The sensitivity of yield to hazard exposure is therefore obtained by differentiation with respect to the *MHI* (Equation 4):

$$\frac{dY_t}{dMHI_t} = \beta_1 + \beta_3 ABI_t \quad (4)$$

In the insurance framework, the tick value represents the monetary change in yield per unit change in the hazard index. In this study, a dynamic tick is therefore defined as (Equation 5):

$$\text{tick} = \frac{\text{dY}}{\text{dMHI}} \times \text{P} = (\beta_1 + \beta_3 \text{ ABI }) \times \text{P} \qquad (5)$$

where *P* denotes the market price of crop. Under this dynamic tick structure, farmers exhibiting higher behavioral adaptation experience lower yield sensitivity to hazard exposure, resulting in smaller expected losses.

3.4 Financial module

The financial module determines the insurance premium using an option-pricing structure in which payouts depend on climate hazard-index thresholds. Because the tick value depends on farmers' behavioral adaptation, the expected payout and resulting premium automatically incorporate behavioral effects. Premium values are computed using a Black-Scholes pricing method (Equations 6-7):

$$\text{Premium}_{\text{call}} = \text{M} \times \text{e}^{-\text{rT}}\, \text{N}(\text{d}_2) \qquad (6)$$

$$\text{Premium}_{\text{put}} = \text{M} \times \text{e}^{-\text{rT}}\, \text{N}(-\text{d}_2) \qquad (7)$$

Where *M* denotes the maximum payout level, *r* represents the risk-free interest rate, and T denotes the duration of the coverage. The parameter $\mathcal{N}(.)$ represents the cumulative distribution function of the standard normal distribution. The term d_2 is defined as (Equation 8):

$$\text{d}_2 = \frac{\text{Ln}\left(\frac{\text{I}_{tr}}{\text{I}_{ex}}\right) + (\text{r} - \frac{\sigma^2}{2})\text{T}}{\sigma\sqrt{\text{T}}} \qquad (8)$$

where I_{tr} represents the trigger *MHI* level, I_{ex} represents the exit *MHI* level, and σ denotes the standard deviation of the normalized *MHI.*

Because the tick value is defined as a function of behavioral adaptation, the expected payout distribution shifts according to the value of ABI_t. Higher levels of adaptive behavior reduce yield sensitivity to hazard exposure, thereby lowering the magnitude of potential losses and, consequently, reducing the actuarially fair premiums.

4 Conclusion

This study introduces the BBII model as a theoretical extension of the multi-hazard index-based insurance model that explicitly endogenizes farmers' behavior within vulnerability estimation and actuarial pricing. By integrating a behavior-based index into the vulnerability module, the framework allows yield sensitivity to climate hazards to vary dynamically with observed behavioral adaptation, which in turn, affects tick values and insurance premiums. This dynamic linkage ensures that insurance pricing reflects not only hazard exposure but also the degree to which farmers respond to early-warning signals through adaptive on-farm actions. Thus, the BBII model addresses a central limitation of conventional index-based insurance models, which treat vulnerability as independent of farmer decision-making processes. In particular, by making yield sensitivity explicitly contingent on adaptive behavior, the BBII model yields a behavior-sensitive premium structure in which effective responses to early-warning signals are associated with lower expected losses and more favorable pricing outcomes.

Methodologically, the proposed framework preserves the core structure of the multi-hazard index-based insurance framework while introducing a behavior-hazard interaction that directly links adaptive responses to pricing outcomes through a dynamic tick mechanism. This approach aligns premium determination with both climate hazards and farmers' behavior, reinforcing the role of early-warning signals and data-driven decision support in agricultural risk management. Rather than changing the index trigger, the BBII model adjusts yield sensitivity and payout magnitude based on farmers' behavior, preserving transparency while allowing premiums and losses to better reflect behavioral differences. The framework is also consistent with ongoing advances in sensor-based monitoring, data analytics, and AIoT-enabled decision-support systems, which increasingly make adaptive farmers' responses observable and measurable in real time.

Future research should focus on the empirical implementation and validation of the BBII framework. Key priorities include developing practical methods to operationalize the ABI using observational data, testing alternative specifications of the behavior-hazard interaction, and evaluating the performance of behavior-sensitive pricing across different farming systems and hazard conditions. Further work is also needed to address institutional and policy issues related to behavioral data governance and sharing, the design of incentives that encourage farmer participation, and coordination between insurance providers and digital decision-support systems issuing early-warning alerts. Addressing these challenges will be essential for translating behavior-based insurance concepts into practical and equitable agricultural risk-management instruments. Overall, the BBII model provides a flexible foundation for advancing dynamic, behavior-aware approaches to agricultural risk management and for strengthening the integration between climate adaptation and risk-transfer mechanisms.

5 Ethics declaration

This paper is theoretical and no ethical clearance is required.

6 AI declaration

Microsoft Copilot was used for language editing only.

References

Abdul-Azeez, O., Ihechere, A. O., & Idemudia, C. (2024). Enhancing business performance: The role of data-driven analytics in strategic decision-making. *International Journal of Management & Entrepreneurship Research*, *6*(7), 2066-2081. https://doi.org/10.51594/ijmer.v6i7.1257

Agnolucci, P., & De Lipsis, V. (2020). Long-run trend in agricultural yield and climatic factors in Europe. *Climatic Change*, *159*(3), 385-405. https://doi.org/10.1007/s10584-019-02622-3

Ardakani, Z., Agnusdei, G. P., Zouari, A., & Miglietta, P. P. (2026). Developing a multi-hazard risk index-based insurance for viticulture: theoretical framework and empirical application. *Agricultural and Food Economics*, *14*(1), 19. https://doi.org/10.1186/s40100-026-00462-1

Angon, P. B., Das, A., Roy, A. R., Khan, J. J., Ahmad, I., Biswas, A., ... & Yeasmin, S. T. (2024). Plant development and heat stress: role of exogenous nutrients and phytohormones in thermotolerance. *Discover Plants*, *1*(1), 17. https://doi.org/10.1007/s44372-024-00020-3

Araújo, S. O., Peres, R. S., Filipe, L., Manta-Costa, A., Lidon, F., Ramalho, J. C., & Barata, J. (2023). Intelligent data-driven decision support for agricultural systems-ID3SAS. *IEEE Access*, *11*, 115798-115815. https://doi.org/10.1109/ACCESS.2023.3324813

Arora, N. K. (2019). Impact of climate change on agriculture production and its sustainable solutions. *Environmental sustainability*, *2*(2), 95-96. https://doi.org/10.1007/s42398-019-00078-w

Aslam, M. A., Ahmed, M., Hassan, F. U., Afzal, O., Mehmood, M. Z., Qadir, G., ... & Hussain, T. (2021). Impact of temperature fluctuations on plant morphological and physiological traits. *In Building climate resilience in agriculture: theory, practice and future perspective* (pp. 25-52). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-79408-8_3

Battaglia, M., Ceglia, I., Calabrese, M., & Iandolo, F. (2025). Systemic risk management and stakeholder engagement: insights from business CSR disclosure. *Corporate Social Responsibility and Environmental Management*, *32*(3), 4295-4314. https://doi.org/10.1002/csr.3186

Basnet, B., & Bang, J. (2018). The state-of-the-art of knowledge-intensive agriculture: a review on applied sensing systems and data analytics. *Journal of Sensors*, *2018*(1), 3528296. https://doi.org/10.1155/2018/3528296

Behera, S., Reddy, N. T. K., & Pradhan, S. (2025). Development of an Intelligent Agricultural Decision Support System for Crop Recommendation Using Machine Learning Techniques. *Next-Generation Computing Systems and Technologies*, *1*(1), 43-53. https://doi.org/10.62762/NGCST.2025.340075

Benfradj, A., Thaljaoui, A., Moulahi, T., Khan, R. U., Alabdulatif, A., & Lorenz, P. (2024). Integration of artificial intelligence (AI) with sensor networks: Trends, challenges, and future directions. *Journal of King Saud University-Computer and Information Sciences*, *36*(1), 101892. https://doi.org/10.1016/j.jksuci.2023.101892

Booth, L., Schueller, L. A., Scolobig, A., & Marx, S. (2020). Stakeholder solutions for building interdisciplinary and international synergies between climate change adaptation and disaster risk reduction. *International journal of disaster risk reduction*, *46*, 101616. https://doi.org/10.1016/j.ijdrr.2020.101616

Carletto, C. (2021). Better data, higher impact: improving agricultural data systems for societal change. *European Review of Agricultural Economics*, *48*(4), 719-740. https://doi.org/10.1093/erae/jbab030

Capone, D., Agnusdei, L., Miglietta, P. P., Zouari, A., & Agnusdei, G. P. (2026). A theoretical framework for vineyard irrigation management using AI-driven sensor systems. *Procedia Computer Science*, *277*, 677-684. https://doi.org/10.1016/j.procs.2026.02.109

Coffey, K., Haile, M., Halperin, M., Wamukoya, G., Hansen, J., Kinyangi, J., & Tesfaye Fantaye, K. (2015). Expanding the contribution of early warning to climate-resilient agricultural development in Africa. https://hdl.handle.net/10568/66596

Constable, A. J., French, S., Karoblyte, V., & Viner, D. (2022). Decision-making for managing climate-related risks: Unpacking the decision process to avoid "trial-and-error" responses. *Frontiers in Climate*, *4*, 754264. https://doi.org/10.3389/fclim.2022.754264

Chen, Z., Galli, M., & Gallavotti, A. (2022). Mechanisms of temperature-regulated growth and thermotolerance in crop species. *Current opinion in plant biology*, *65*, 102134. https://doi.org/10.1016/j.pbi.2021.102134

Chen, T., Zhang, L., Wen, M., Yuan, W., & Lin, W. (2025). Can the development of agricultural insurance promote the resilience of agricultural economy? The dynamic mechanisms of the digital economy development. *International Review of Economics & Finance*, *103*, 104386. https://doi.org/10.1016/j.iref.2025.104386

Chowdhury, S., Sayem, N. S., Ali, M. R., Haque, A. O., Alam, M. S., Saha, C. K., & Rahman, K. S. (2025). Enhanced IoT-Based Smart Agro-Farm Security System Through Multi-Sensor Fusion and Field Validation. *Smart Agricultural Technology*, 101512. https://doi.org/10.1016/j.atech.2025.101512

Devore, J. (2007). Making sense of data: A practical guide to exploratory data analysis and data mining. https://doi.org/10.1198/tas.2007.s144

Di Gangi, M., Belcore, O. M., & Polimeni, A. (2023). An overview on decision support systems for risk management in emergency conditions: present, past and future trends. *International journal of transport development and integration*, *7*(1), 45-53. https://doi.org/10.18280/ijtdi.070106

Fang, H., Liu, F., Nie, P., & Wu, J. (2021). Agricultural decision-making methods and systems. *In Agricultural Internet of Things: Technologies and Applications* (pp. 251-290). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-65702-4_9

Gamage, A., Gangahagedara, R., Subasinghe, S., Gamage, J., Guruge, C., Senaratne, S., ... & Merah, O. (2024). Advancing sustainability: The impact of emerging technologies in agriculture. *Current Plant Biology*, *40*, 100420. https://doi.org/10.1016/j.cpb.2024.100420

Gesualdo, G. C., Benso, M. R., Sass, K. S., & Mendiondo, E. M. (2024). Index-based insurance to mitigate current and future extreme events financial losses for water utilities. *International Journal of Disaster Risk Reduction*, *100*, 104218. https://doi.org/10.1016/j.ijdrr.2023.104218

Goud, K. S., Jansi, Y., Kumar, R. A., Patnala, E., & Jyothi, R. S. (2024, April). IADSS: Integrated Agricultural Decision Support System Using Machine Learning. In *International Conference on Cognitive Computing and Cyber Physical Systems* (pp. 378-391). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-77081-4_30

Gudivada, V. N. (2025). Data analytics: fundamentals. In *Data analytics for intelligent transportation systems* (pp. 27-66). Elsevier. https://doi.org/10.1016/B978-0-12-809715-1.00002-X

Hou, D., & Wang, X. (2025). Can agricultural insurance enhance comprehensive grain production capacity? Mechanisms of risk protection and production incentives. *Frontiers in Sustainable Food Systems*, *9*, 1649495. https://doi.org/10.3389/fsufs.2025.1649495

Janni, M., Maestri, E., Gullì, M., Marmiroli, M., & Marmiroli, N. (2024). Plant responses to climate change, how global warming may impact on food security: a critical review. *Frontiers in plant science*, *14*, 1297569. https://doi.org/10.3389/fpls.2023.1297569

Kumar, R., Lal, M. K., Tiwari, R. K., Kumar, A., Behera, B., Nayak, L., ... & Altaf, M. A. (2023). Traditional and emerging climate-resilient agricultural practices for enhancing food production and nutritional quality. In *New Frontiers in Plant-Environment Interactions: Innovative Technologies and Developments* (pp. 551-570). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-43729-8_21

Łągiewska, M., & Panek-Chwastyk, E. (2025). Integrating remote sensing and autonomous robotics in precision agriculture: Current applications and workflow challenges. *Agronomy*, *15*(10), 2314. https://doi.org/10.3390/agronomy15102314

Leal, A. C. (2024). Data analytics in agriculture. In *Digital Agriculture: A Solution for Sustainable Food and Nutritional Security* (pp. 519-539). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-031-43548-5_17

Li, H. J., Luo, X. G., Zhang, Z. L., Huang, S. W., & Jiang, W. (2025). A usage-based insurance (UBI) pricing model considering customer retention. Insurance: Mathematics and Economics, 124, 103132. https://doi.org/10.1016/j.insmatheco.2025.103132

Liakos, K. G., Busato, P., Moshou, D., Pearson, S., & Bochtis, D. (2018). Machine learning in agriculture: A review. *Sensors*, *18*(8), 2674. https://doi.org/10.3390/s18082674

MacPherson, J., Voglhuber-Slavinsky, A., Olbrisch, M., Schöbel, P., Dönitz, E., Mouratiadou, I., & Helming, K. (2022). Future agricultural systems and the role of digitalization for achieving sustainability goals. A review: Future agricultural systems and the role of digitalization for achieving sustainability goals. A review. *Agronomy for sustainable development*, *42*(4), 70. https://doi.org/10.1007/s13593-022-00792-6

Manono, B. O., Mwami, B., Mutavi, S., & Nzilu, F. (2026). Precision farming with smart sensors: Current state, challenges and future outlook. *Sensors*, *26*(3), 882. https://doi.org/10.3390/s26030882

Mansoor, S., Iqbal, S., Popescu, S. M., Kim, S. L., Chung, Y. S., & Baek, J. H. (2025). Integration of smart sensors and IOT in precision agriculture: trends, challenges and future prospectives. *Frontiers in Plant Science*, *16*, 1587869. https://doi.org/10.3389/fpls.2025.1587869

Menalled, F. D., Ebel, R., Peterson, R. K., Schutte, B. J., Shergill, L. S., & Smith, R. G. (2026). Eco-evolutionary and risk assessment frameworks for assessing agroecosystem resilience in a changing climate. *Discover Agriculture*, *4*(1), 17. https://doi.org/10.1007/s44279-026-00491-w

Moore, C. E., Meacham-Hensold, K., Lemonnier, P., Slattery, R. A., Benjamin, C., Bernacchi, C. J., ... & Cavanagh, A. P. (2021). The effect of increasing temperature on crop photosynthesis: from enzymes to ecosystems. *Journal of experimental botany*, *72*(8), 2822-2844. https://doi.org/10.1093/jxb/erab090

Mora, M., Phillips-Wren, G., & Wang, F. (2018). The role of decision-making support systems in risk management. *In The Routledge companion to risk, crisis and security in business* (pp. 441-456). Routledge. https://doi.org/10.4324/9781315629520

Minoli, S., Jägermeyr, J., Asseng, S., Urfels, A., & Müller, C. (2022). Global crop yields can be lifted by timely adaptation of growing periods to climate change. *Nature Communications*, *13*(1), 7079. https://doi.org/10.1038/s41467-022-34411-5

Mishra, H., & Mishra, D. (2024). AI for data-driven decision-making in smart agriculture: from field to farm management. In *Artificial intelligence techniques in smart agriculture* (pp. 173-193). Singapore: Springer Nature Singapore. https://doi.org/10.1007/978-981-97-5878-4_11

Muffler, L., Weigel, R., Beil, I., Leuschner, C., Schmeddes, J., & Kreyling, J. (2024). Winter and spring frost events delay leaf-out, hamper growth and increase mortality in European beech seedlings, with weaker effects of subsequent frosts. *Ecology and Evolution*, *14*(7), e70028. https://doi.org/10.1002/ece3.70028

Nakalembe, C., Becker-Reshef, I., Bonifacio, R., Hu, G., Humber, M. L., Justice, C. J., ... & Sanchez, A. (2021). A review of satellite-based global agricultural monitoring systems available for Africa. *Global food security*, *29*, 100543. https://doi.org/10.1016/j.gfs.2021.100543

Padhiary, M., Kumar, A., & Sethi, L. N. (2025). Emerging technologies for smart and sustainable precision agriculture. *Discover Robotics*, *1*(1), 6. https://doi.org/10.1007/s44430-025-00006-0

Palutikof, J. P., Street, R. B., & Gardiner, E. P. (2019). Decision support platforms for climate change adaptation: an overview and introduction. *Climatic Change*, *153*(4), 459-476. https://doi.org/10.1007/s10584-019-02445-2

Parnell, R. A., & Gangwish, A. (2023). Effective collaborative decision-making includes stakeholder analysis and communication. *In A Practical Guide for Developing Cross-Disciplinary Collaboration Skills* (pp. 103-123). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-031-37220-9_6

Power, D. J. (2008). Understanding data-driven decision support systems. *Information Systems Management*, *25*(2), 149-154. https://doi.org/10.1080/10580530801941124

Power, D. J. (2016). Data science: supporting decision-making. *Journal of Decision systems*, *25*(4), 345-356. https://doi.org/10.1080/12460125.2016.1171610

Raza, A., Razzaq, A., Mehmood, S. S., Zou, X., Zhang, X., Lv, Y., & Xu, J. (2019). Impact of climate change on crops adaptation and strategies to tackle its outcome: *A review. Plants*, *8*(2), 34. https://doi.org/10.3390/plants8020034

Rejeb, A., Rejeb, K., & Hassoun, A. (2025). The impact of machine learning applications in agricultural supply chain: a topic modeling-based review. *Discover Food*, *5*(1), 141. https://doi.org/10.1007/s44187-025-00419-1

Ricart, S., Gandolfi, C., & Castelletti, A. (2023). Climate change awareness, perceived impacts, and adaptation from farmers' experience and behavior: a triple-loop review. *Regional Environmental Change*, *23*(3), 82. https://doi.org/10.1007/s10113-023-02078-3

Rinaldi, M., & He, Z. (2014). Decision support systems to manage irrigation in agriculture. *Advances in agronomy*, *123*, 229-279. https://doi.org/10.1016/B978-0-12-420225-2.00006-6

Saleem, A., Anwar, S., Nawaz, T., Fahad, S., Saud, S., Ur Rahman, T., ... & Nawaz, T. (2025). Securing a sustainable future: the climate change threat to agriculture, food security, and sustainable development goals. *Journal of Umm Al-Qura University for Applied Sciences*, *11*(3), 595-611. https://doi.org/10.1007/s43994-024-00177-3

Sarker, I. H. (2021). Data science and analytics: an overview from data-driven smart computing, decision-making and applications perspective. SN *Computer Science*, *2*(5), 377. https://doi.org/10.1007/s42979-021-00765-8

Savari, M., Damaneh, H. E., Damaneh, H. E., & Cotton, M. (2023). Integrating the norm activation model and theory of planned behaviour to investigate farmer pro-environmental behavioural intention. *Scientific Reports*, *13*(1), 5584. https://doi.org/10.1038/s41598-023-32831-x

Selvaraju, R. (2012). Climate risk assessment and management in agriculture. *Building resilience for adaptation to climate change in the agriculture sector*, *23*(71), 20. https://openknowledge.fao.org/.pdf

Shah, I. H., Manzoor, M. A., Jinhui, W., Li, X., Hameed, M. K., Rehaman, A., ... & Chang, L. (2024). Comprehensive review: Effects of climate change and greenhouse gases emission relevance to environmental stress on horticultural crops and management. *Journal of Environmental Management*, *351*, 119978. https://doi.org/10.1016/j.jenvman.2023.119978

Song, H., & Zhu, Z. (2025). Farmers' adaptive behaviors to climate change and their influencing factors: evidence from the Guanzhong Region of Shaanxi, China. *Frontiers in Sustainable Food Systems*, *9*, 1648301. https://doi.org/10.3389/fsufs.2025.1648301

Soussi, A., Zero, E., Sacile, R., Trinchero, D., & Fossa, M. (2024). Smart sensors and smart data for precision agriculture: a review. *Sensors*, *24*(8), 2647. https://doi.org/10.3390/s24082647

Singh, P., & Krishnamurthi, R. (2024). Empowering agricultural security with IoT and deep learning driven real-time alert systems. *International Journal of Information Technology*, *16*(7), 4627-4635. https://doi.org/10.1007/s41870-024-02074-8

Tselentis, D. I., Yannis, G., & Vlahogianni, E. I. (2016). Innovative insurance schemes: pay as/how

you drive. Transportation Research Procedia, 14, 362-371. https://doi.org/10.1016/j.trpro.2016.05.088

Vandana, P., Gupta, A., & Kumar, M. (2024). Drought and waterlogging stress responses in crops. *In Plant-Microbe Interaction and Stress Management* (pp. 51-78). Singapore: Springer Nature Singapore. https://doi.org/10.1007/978-981-97-4239-4_4

Waqas, M., Naseem, A., Humphries, U. W., Hlaing, P. T., Dechpichai, P., & Wangwongchai, A. (2025). Applications of machine learning and deep learning in agriculture: A comprehensive review. *Green Technologies and Sustainability*, *3*(3), 100199. https://doi.org/10.1016/j.grets.2025.100199

Xing, Y., Liu, X., & Wang, X. (2025). Integrating UAVs, satellite remote sensing, and machine learning in precision agriculture: pathways to sustainable food production, resource efficiency, and scalable innovation. *Frontiers in Agronomy*, *7*, 1670380. https://doi.org/10.3389/fagro.2025.1670380

Xu, H., Lu, Y., Chen, Y., & Fan, Y. (2026). Warning response behavior model: conceptual clarification and operationalization. *Natural Hazards*, *122*(5), 200. https://doi.org/10.1007/s11069-026-08003-1

Zhai, Z., Martínez, J. F., Beltran, V., & Martínez, N. L. (2020). Decision support systems for agriculture 4.0: Survey and challenges. *Computers and Electronics in Agriculture*, *170*, 105256. https://doi.org/10.1016/j.compag.2020.105256

Zhang, Y., Chen, X., Geng, S., & Zhang, X. (2025). A review of soil waterlogging impacts, mechanisms, and adaptive strategies. *Frontiers in Plant Science*, *16*, 1545912. https://doi.org/10.3389/fpls.2025.1545912

Zhu, W., Zhang, J., Porth, L., & Tan, K. S. (2025). Index-based insurance design for climate and weather risk management: a review. *Quantitative Risk Management in Agricultural Business*, 5-39. https://doi.org/10.1007/978-3-031-80574-5_2

The Impact of Knowledge Asymmetries on Business–IT Alignment Dynamics

Fabrizio Amarilli
Dublin City University, Dublin, Ireland

Abstract

Shared Domain Knowledge (SDK) between business and IT executives is a crucial antecedent of business-IT alignment success. However, existing research conceptualizes SDK as a unidimensional attribute varying primarily in magnitude, assuming increasing shared knowledge uniformly improves alignment. This aggregated view cannot explain why organizations with similar SDK levels experience markedly different alignment trajectories. This study investigates how knowledge asymmetry configurations affect business-IT alignment by decomposing SDK into a multidimensional profile. Drawing on knowledge management theory, we develop a framework specifying four core knowledge dimensions: technical-strategic IT knowledge distribution, feasibility-possibility knowledge calibration, boundary-spanning knowledge, and temporal-relational knowledge. These dimensions interact to create distinct patterns of knowledge asymmetry with predictable alignment dynamics. We conducted a longitudinal comparative case study of three manufacturing companies implementing enterprise systems over 15 to 24 months. Data included semi-structured interviews with CEOs, CIOs, and key stakeholders, supplemented by project documentation and assessment reports. Our analysis identified three patterns. Bridging Asymmetry occurs when extreme knowledge gaps are compensated by one executive's dual competencies enabling effective translation, producing successful alignment despite persistent asymmetry. Complementary Asymmetry emerges when neither party possesses boundary-spanning capability, requiring external intervention for partial alignment. Aspirational Asymmetry arises when both parties share high strategic IT knowledge but diverge in feasibility understanding, creating expectation-reality gaps leading to knowledge atrophy. Our findings make three contributions. First, we demonstrate that knowledge configuration matters more than aggregate level in determining alignment outcomes. Second, we identify boundary-spanning knowledge as the critical mechanism through which asymmetries either enable or constrain alignment. Third, we reveal that high SDK can generate problems through unrealistic expectations and success-induced complacency, challenging the assumption that more shared knowledge invariably improves outcomes. These findings advance theoretical understanding of alignment's social dimension and offer practitioners diagnostic tools for managing knowledge asymmetries effectively.

Keywords: Business IT alignment, Social alignment, Shared domain knowledge, Knowledge management.

Paper type: Academic Research Paper.

1 Introduction

Business-IT alignment remains a perennial concern for researchers and practitioners (Henderson and Venkatraman, 1993; Luftman et al., 2017). Over three decades of research, scholars have identified multiple dimensions through which alignment manifests. Among these, the social dimension has received particular attention for enabling mutual understanding between business and IT executives (Reich and Benbasat, 1996, 2000). This understanding is operationalized as Shared Domain Knowledge (SDK), defined as the degree to which business and IT executives understand and can participate in each other's domains (Reich and Benbasat, 2000; Preston and Karahanna, 2009). While existing research treats SDK as a crucial alignment antecedent, most studies conceptualize it as a single attribute varying from low to high (Johnson and Lederer, 2010). We challenge this monolithic view, proposing instead that SDK constitutes a multidimensional knowledge profile whose configuration shapes alignment dynamics in distinct ways.

The dominant conceptualization assumes increasing shared knowledge improves alignment (Reich and Benbasat, 2000; Preston and Karahanna, 2009), yet this aggregated view obscures critical variation. Several research strands hint at SDK's multidimensional character. Studies of CIO capabilities distinguish technical from strategic IT knowledge (Peppard et al., 2011; Chun and Mooney, 2009). Research on IT project failures identifies gaps between understanding technological possibilities versus implementation feasibility (Keil et al., 2000; Lyytinen and Robey, 1999). Work on cross-functional collaboration emphasizes that domain expertise alone is insufficient without boundary-spanning mechanisms (Carlile, 2004; Levina and Vaast, 2005). Despite these insights, no framework systematically integrates these dimensions or examines how their configuration affects alignment. Drawing on knowledge management theory (Carlile, 2004; Levina and Vaast, 2005), we decompose SDK into four dimensions: technical-strategic IT knowledge distribution, feasibility-possibility knowledge calibration, boundary-spanning knowledge, and temporal-relational knowledge (see Table 2 for full

operationalization). These dimensions do not contribute additively to alignment. Rather, their specific configuration creates distinct patterns of knowledge asymmetry with different alignment dynamics. Our research question asks: How do different configurations of knowledge asymmetries affect the business-IT alignment process?

We conducted a longitudinal comparative case study of three manufacturing companies implementing enterprise systems over 15 to 24 months. Our analysis identified three distinct patterns. Bridging Asymmetry occurs when extreme knowledge gaps are compensated by one party's dual competencies enabling effective translation. Complementary Asymmetry emerges when neither party possesses boundary-spanning capability, requiring external intervention. Aspirational Asymmetry arises when both parties share high strategic IT knowledge but diverge in feasibility understanding, creating expectation-reality gaps leading to knowledge atrophy.

This study makes three contributions. First, we demonstrate that knowledge configuration matters more than aggregate level. Second, we identify boundary-spanning knowledge as a critical mechanism through which asymmetries either enable or constrain alignment. Third, we challenge the assumption that higher SDK invariably improves alignment, showing how high strategic IT knowledge can generate problems through unrealistic expectations and success-induced complacency.

2 Literature Review

Business-IT alignment refers to the degree of congruence between business objectives and IT capabilities, encompassing both the process through which this fit is achieved and the state of alignment itself (Chan and Reich, 2007). Over three decades, scholars have proposed multiple dimensions through which alignment manifests (Henderson and Venkatraman, 1993; Reich and Benbasat, 2000). Reich and Benbasat (2000) distinguish between intellectual alignment, concerning the fit between business and IT strategy, and social alignment, addressing the state in which executives understand and are committed to business and IT missions, objectives, and plans. Subsequent research expanded this to include operational alignment, referring to congruence between business and IT processes and structures, and strategic alignment at the organizational level (Gerow et al., 2014). Social alignment has received sustained attention as an antecedent condition. Gerow et al.'s (2014) meta-analytic review demonstrates that social alignment influences both intellectual and operational alignment, which affect organizational performance, positioning it as foundational.

Within social alignment, Shared Domain Knowledge serves as a central component (Reich and Benbasat, 2000; Preston and Karahanna, 2009). SDK refers to executives' ability to understand and participate effectively in each other's key processes (Reich and Benbasat, 2000, p.82). Preston and Karahanna (2009) identify SDK as one of four antecedents of IS strategic alignment. Johnson and Lederer (2010) demonstrate that CEO/CIO mutual understanding correlates positively with strategic alignment and IS contribution to organizational outcomes. These studies establish SDK as an empirically validated predictor of alignment success, yet conceptualize it as a unidimensional construct varying primarily in magnitude.

This aggregated view overlooks substantive variation in knowledge composition. Research on CIO capabilities suggests different knowledge types produce different outcomes. Peppard et al. (2011) distinguish technical, business, and interpersonal competence, noting that effective IT leadership requires mastery across multiple domains. Chun and Mooney (2009) document a shift from technical management toward strategic partnership reflecting changing knowledge requirements. Bassellier et al. (2001) develop an IT competence measure encompassing explicit and tacit knowledge in technical and managerial dimensions, demonstrating that knowledge types cannot be reduced to a single scale. Armstrong and Sambamurthy (1999) distinguish between knowledge of IT's strategic potential and technical implementation requirements. These studies indicate that knowledge nature matters, not merely quantity.

A second research stream examines discrepancies between understanding technological possibilities and operational realities. Ewusi-Mensah (1997) identifies the gap between initial expectations and feasibility assessments as a critical failure factor. Keil et al. (2000) find that decision-makers often persist with failing initiatives because their understanding of what technology should deliver obscures recognition of what it can achieve under prevailing constraints. Lyytinen and Robey (1999) argue that organizations frequently fail to adjust expectations when confronted with infeasibility evidence, suggesting systematic difficulty in calibrating possibility against feasibility knowledge.

Knowledge management scholarship provides theoretical resources for understanding how distributed expertise is integrated across boundaries. Carlile (2004) demonstrates that knowledge integration depends on actors' capacity to recognize dependencies, establish common meanings, and negotiate conflicting interests. Levina and Vaast (2005) identify boundary spanning competence as an emergent practice requiring ongoing

relational work. Majchrzak et al. (2012) reveal that performance depends on actors who can integrate dispersed expertise. Pawlowski and Robey (2004) find that successful integration of business and technical knowledge requires individuals who understand both domains and can translate between them. This literature establishes that knowledge overlap alone does not guarantee effective collaboration; boundary-spanning mechanisms are essential.

Temporal aspects of knowledge have received less systematic attention despite their documented influence. Orlikowski and Yates (2002) argue that timing, rhythm, and duration shape how work unfolds and technologies are adopted. Sabherwal and Robey (1995) note that alignment must be understood as a temporal phenomenon rather than a static state. Brown and Grant (2005) observe that misalignments between business planning cycles and IT development timelines create persistent coordination challenges. Xue et al. (2008) demonstrate that temporal characteristics of IT investments moderate the relationship between governance structures and decision outcomes.

The literature reveals multiple knowledge dimensions that existing SDK frameworks do not explicitly theorize. Technical knowledge differs from strategic knowledge; understanding possibilities differs from understanding feasibility; domain expertise differs from boundary-spanning capability; and awareness of temporal dynamics shapes how knowledge is applied. Current conceptualizations cannot account for configurations where executives possess high knowledge in one dimension but low in another, nor explain why such configurations produce systematically different outcomes. A multidimensional framework specifying how knowledge types interact offers potential to resolve these gaps.

3 Methodology and Theoretical Framework

This study employs a multiple case study approach to investigate how knowledge asymmetry configurations affect business-IT alignment. Three organizations undertaking alignment initiatives through enterprise system implementations served as research sites over 15 to 24 months, enabling analysis of how knowledge configurations evolved and shaped alignment dynamics.

Case study methodology is appropriate for investigating complex social processes embedded in organizational contexts (Eisenhardt, 1989; Yin, 2014). Our research question asks "how" knowledge configurations affect alignment, requiring a method that accommodates emergent findings. The multidimensional nature of SDK demands rich contextual data capturing interplay between knowledge dimensions and organizational dynamics (Langley, 1999).

The three cases provide theoretical sampling across contexts and outcomes. Company 1 is a family-owned sports nutrition manufacturer (200 employees, €52M revenue) that hired a new CIO and undertook transformation following an unsuccessful ERP implementation. Company 2 operates in climate solutions (200 employees, €42M revenue), implementing enterprise applications following external assessment. Company 3 is a large European manufacturer that launched a 24-month alignment initiative implementing a comprehensive enterprise system.

Data collection followed established protocols (Miles and Huberman, 1994). Primary data came from semi-structured interviews with key informants. Interviews lasted 60-90 minutes and were recorded and transcribed where permitted; field notes supplemented non-recorded conversations. Secondary data included project documentation, steering committee minutes, external assessment reports, and internal communications, enabling triangulation and providing temporal markers.

Table 1: Data Sources

Case	Interviews	Informants	Secondary Sources	Time Period
Company 1	6 interviews	President, CEO, CIO (3 meetings), CFO, Head of Operations, Head of Sales	Project documentation, external assessment report, financial communications	2019-2020
Company 2	8 interviews	CEO, CFO, IT Manager/CIO, IT Staff, Head of Production, Head of R&D, Sales/Marketing Manager, Head of Post-Sales	Project documentation, steering committee minutes, external assessment report (Fondazione Politecnico di Milano), supplier presentations	2018-2019
Company 3	11 interviews	CEO, CIO, External consultants, Project team members	Steering committee documentation, project reports, university assessment reports	2019-2021

Data analysis proceeded through iterative coding and pattern recognition (Strauss and Corbin, 1998). Initial open coding identified knowledge gaps, translation difficulties, expectation mismatches, and coordination challenges. These codes were organized into higher-order categories corresponding to four knowledge dimensions emerging from knowledge management and alignment literatures. For example, when the Company 2 CEO stated "*I don't understand anything; there is no reason for me to stay here*," this was coded as "low technical IT knowledge." The CIO's statement that the system represented "nothing special" was coded as "inability to articulate strategic value." The CFO's need for "*someone with credibility to say this guy is not saying nonsense*" revealed absent boundary-spanning capability. Patterns emerged when similar configurations appeared across cases with similar outcomes.

The analytical framework specifies four knowledge dimensions derived from research on CIO capabilities, knowledge boundaries, and organizational learning.

Table 2: Knowledge Dimensions Framework

Dimension	Sub-Dimensions	What It Measures	Key References
Technical-Strategic IT Knowledge Distribution	Technical IT knowledge; Strategic IT knowledge; Business domain knowledge	How IT knowledge (technical vs strategic) and business knowledge are distributed between CEO and CIO	Bassellier et al., 2001; Peppard et al., 2011; Chun and Mooney, 2009
Feasibility-Possibility Knowledge Calibration	Possibility knowledge; Feasibility knowledge; Expectation gaps	Whether executives share realistic understanding of achievable outcomes versus theoretical capabilities	Keil et al., 2000; Lyytinen and Robey, 1999; Ewusi-Mensah, 1997
Boundary-Spanning Knowledge	Convergent knowledge; Complementary knowledge; Translation capability	Whether knowledge overlaps or requires integration, and who possesses bridging capability	Carlile, 2004; Levina and Vaast, 2005; Pawlowski and Robey, 2004; Majchrzak et al., 2012
Temporal-Relational Knowledge	Temporal alignment; Mutual knowledge; Supplier ecosystem knowledge	Understanding of time horizons, pacing, mutual constraints, and supplier dynamics	Orlikowski and Yates, 2002; Sabherwal and Robey, 1995; Brown and Grant, 2005

This framework guided data collection and analysis, enabling systematic comparison while remaining open to emergent patterns.

4 Results

Our analysis of the three cases revealed distinct patterns in how knowledge asymmetries shape alignment processes. Rather than organizing findings by individual case, we present results according to three emergent patterns identified through systematic cross-case comparison. Each pattern is characterized by a specific configuration of knowledge dimensions and produces predictable alignment dynamics. The patterns demonstrate that knowledge distribution and bridging capability matter more than aggregate knowledge levels in determining alignment outcomes. Table 3 provides a cross-case comparison across the four knowledge dimensions, revealing the configurations that differentiate the three patterns.

Table 3: Cross-Case Comparison

Dimension	Bridging	Complementary	Aspirational
CEO IT Knowledge	Zero technical; very low strategic	Low technical; low-medium strategic ("IT faith" not knowledge)	Low-medium technical; high strategic
CIO IT Knowledge	High technical; high strategic	High technical; low-medium strategic	High technical; high strategic
CIO Business Knowledge	HIGH (consulting background; "*speaks CEO's language*")	LOW (technical focus; cannot translate value)	High (business-savvy; understands operations)
Feasibility-Possibility Gap	Bridged by CIO translation	CEO questions investment itself (backward-looking)	High expectations on both sides create reality gap
Boundary-Spanning	CIO exclusively (personal capability)	NOBODY internally (external actors required)	Both initially (steering committee); plateaued
Temporal Alignment	CIO manages (adapts to seasonality)	Misaligned (CEO backward-looking; budget vicious circle)	Initially aligned; drifted (internal vs. external tempo)

Dimension	Bridging	Complementary	Aspirational
Key Mechanism	Personal translation substitutes for convergence	External compensation for double gap; demonstration over articulation	Expectation-reality gap; knowledge atrophy from success-induced complacency
Alignment Outcome	Success despite asymmetry	Partial success via external support	Initial success, then drift
Fragility Source	Individual dependence	Externally sourced capability	Knowledge plateau in dynamic environment

4.1 Pattern 1: Bridging asymmetry

The first pattern emerges when extreme knowledge asymmetry is compensated by one executive's dual competencies enabling translation. Case 1 exemplifies this through a critical feature absent in the other cases. The CEO possessed limited IT knowledge ("*zero IT skills*" and "*very limited commitment to IT*") and had experienced a failed ERP project. By contrast, the newly hired CIO brought deep technical expertise and consulting experience that equipped him with business domain knowledge. The critical capability distinguishing Case 1 from Cases 2 and 3 was that the CIO "*could understand and speak the language of the two managers/owners*."

This boundary-spanning capability produced successful alignment despite persistent asymmetry. The CEO never developed IT literacy, unlike Case 3 where both executives developed shared understanding. Instead, the CIO performed continuous translation, a capability absent in Case 2 where the CIO lacked business knowledge. When confronting supplier dependencies, the CIO presented choices in business terms owners could evaluate, such as paying the IT supplier versus hiring internal developers. Trust became the enabling mechanism: "*they trusted him and developed trust in IT*."

Cross-case comparison reveals why this pattern succeeded where others struggled. Unlike Case 2's double gap, Case 1's CIO could translate in both directions. Unlike Case 3's eventual plateau, Case 1 sustained alignment through continuous translation. However, this pattern exhibits structural fragility. The observation that owners "thought I would do everything" reveals absent institutional structure to sustain alignment should the CIO depart.

4.2 Pattern 2: Complementary asymmetry

The second pattern materializes when neither party possesses boundary-spanning knowledge, creating a double gap. Case 2 provides the clearest instantiation. The CEO demonstrated low technical IT knowledge, leaving a technical presentation with the explanation "*I don't understand anything; there is no reason for me to stay here*." More significantly, his strategic IT knowledge took the form of "*strategic IT faith*" rather than understanding, a distinction not observed in Cases 1 or 3. He approved a relevant investment and established governance structures, yet his reasoning revealed a backward-looking temporal frame. When confronted with the investment proposal, he asked "*we used this information system for more than 20 years and our market position is still good... what is the performance that will be improved*?" This inverted competitive logic contrasts sharply with Case 3's forward-looking awareness.

The CIO profile presented a mirror-image deficit. While technically proficient, the Company 2 CIO lacked the business domain knowledge that enabled Case 1's successful translation. His budget proposals failed repeatedly, and when articulating the ERP system's value, he offered only that it represented "*nothing special*, *we put in place what we should have done before*." This inability to connect technical capability to business outcomes created a "*budget vicious circle*" where each failed translation reduced trust, which constrained resources, which further degraded articulation capability. No such dynamic emerged in Case 1, where business fluency prevented communication breakdowns, or Case 3, where both parties shared sufficient strategic IT knowledge.

Partial alignment required sustained external intervention, a dependency absent in Cases 1 and 3. The University assessment team and IT suppliers performed the boundary-spanning function no internal actor could provide. The CFO's statement captured this: "*I need someone who has the credibility*, *competence and courage to say in front of me that this guy [the CIO] is not saying nonsense*." The CEO's inability to evaluate competing claims created political vulnerability, as when the Head of Production claimed "*the old system has much less bugs*" to justify resistance.

The Company 2 CIO developed "*demonstration over articulation*," pursuing bottom-up operational evidence when value arguments failed. This compensating mechanism succeeded in commercial functions where value was directly observable but failed in production where resistance could be sustained through appeals to the CEO's technical uncertainty.

4.3 Pattern 3: Aspirational asymmetry

The third pattern arises when both parties possess high strategic IT knowledge but diverge in understanding what is feasibly achievable, a configuration absent in Cases 1 and 2 where knowledge levels differed substantially. Case 3 manifested this pattern across 24 months. Both CEO and CIO demonstrated sophisticated awareness of IT's transformative potential, contrasting with Case 1's zero CEO knowledge and Case 2's "strategic IT faith" without understanding. The steering committee enabled collaborative knowledge development, leading to the observation that "we are speaking the same language," an achievement not attempted in Case 1 and not achieved in Case 2.

The alignment problem emerged not from knowledge deficits but from knowledge calibration failures. High possibility knowledge on both sides generated aspirational expectations that feasibility understanding could not match. As the deployed system stabilized, knowledge development plateaued. The organization's knowledge became past-oriented, focused on achievements rather than emerging competitive requirements. One executive's concern captured this temporal drift: "technology and competitors are moving fast and I suspect that one day we will realize that we are lagging behind." This gradual erosion differs fundamentally from Case 1's sustained individual-dependent alignment and Case 2's chronic translation failures.

The knowledge atrophy mechanism operated through success-induced complacency. The project's successful completion created a sense that alignment had been "solved," obviating continued knowledge investment. Simultaneously, the internal tempo shifted toward maintenance while external competitive tempo accelerated. Supplier ecosystem knowledge developed earlier, when the organization learned "we will never be their priority" with large vendors, was not sustained. This temporal drift distinguishes Case 3 from Case 1, where the CIO continuously adapted, and Case 2, where temporal misalignment existed from the beginning.

Cross-case comparison reveals that high shared domain knowledge does not immunize organizations against alignment drift. When knowledge convergence plateaus and temporal orientation shifts from future to past, even sophisticated leadership teams can lose alignment with evolving requirements. The mechanism differs fundamentally from Cases 1 and 2, where the problem was insufficient knowledge or absent bridging capability rather than knowledge becoming static.

4.4 Cross-pattern synthesis

Table 4 synthesizes the defining characteristics of each pattern, highlighting the specific knowledge configurations and mechanisms that differentiate them.

Table 4: Summary of Knowledge Asymmetry Patterns

Pattern	Knowledge Configuration	Critical Dimension	Alignment Dynamic	Key Quote	Risk Factor
Bridging Asymmetry	Extreme CEO-CIO gap + CIO has dual competencies (IT + business)	Boundary-Spanning	CIO translates effectively; CEO delegates without understanding	"*He could speak the language of the entrepreneur*"	Over-dependence on individual; if CIO leaves, bridging disappears
Complementary Asymmetry	CEO has "strategic IT faith"; CIO has technical skills but no business knowledge	Boundary-Spanning + Technical-Strategic Distribution	Double gap requires external boundary-spanners; demonstration over articulation	"*I need someone who has credibility to say this guy is not saying nonsense*"	Alignment capability externally sourced; non-transferable
Aspirational Asymmetry	Both have high strategic IT knowledge but feasibility understanding diverges	Feasibility-Possibility + Temporal-Relational	High expectations create reality gaps; knowledge becomes past-oriented	"*Technology and competitors are moving fast and I suspect we are lagging behind*"	Success-induced complacency; static knowledge in dynamic environment

The cases' patterns show that knowledge asymmetries produce distinct and predictable alignment dynamics. The configuration of knowledge across dimensions matters more than aggregate knowledge levels. Bridging capability emerges as the critical differentiator between successful alignment despite asymmetry (Case 1) and coordination failures requiring external compensation (Case 2). High strategic knowledge shared between executives (Case 3) produces different challenges from those observed when knowledge levels differ

substantially (Cases 1 and 2). These findings challenge the prevailing assumption in SDK literature that increasing shared knowledge uniformly improves alignment

5 Discussion and conclusion

Through longitudinal analysis of three cases, we identified three knowledge asymmetry patterns that shape alignment dynamics. Our findings offer theoretical and practical contributions.

5.1 Theoretical Contributions

Our primary contribution lies in decomposing SDK from a unidimensional attribute into a multidimensional profile. Existing research assumes increasing shared knowledge improves alignment (Reich and Benbasat, 2000; Preston and Karahanna, 2009). We demonstrate that this aggregated view obscures critical variation. Organizations with similar aggregate SDK levels experience markedly different trajectories depending on knowledge distribution across technical, strategic, boundary-spanning, and temporal dimensions. Case 1 shows successful alignment despite zero CEO IT knowledge because the CIO's business knowledge enabled translation. Case 2 shows struggles despite moderate knowledge levels because neither party could bridge domains. Configuration, not quantity, determines outcomes.

Second, we identify boundary-spanning knowledge as a critical dimension largely absent from existing SDK frameworks. Traditional research focuses on what executives know but neglects whether anyone can integrate knowledge across boundaries (Carlile, 2004; Levina and Vaast, 2005). Our findings reveal that boundary-spanning capability functions as the mechanism through which asymmetries either enable or constrain alignment. Internal bridging (Case 1) produces different dynamics than its absence (Case 2), even when knowledge gaps appear similar. Alignment research should assess not only knowledge levels but also translation mechanisms.

Third, our findings challenge the assumption that higher SDK invariably improves alignment. The aspirational asymmetry pattern shows how high strategic IT knowledge can generate problems through unrealistic expectations and success-induced complacency. Case 3's knowledge atrophy demonstrates that convergent knowledge can plateau when organizations conclude alignment has been permanently achieved. Knowledge management in alignment contexts requires ongoing investment rather than one-time achievement.

5.2 Practical Contributions

Our framework offers diagnostic value. Organizations can assess their knowledge configuration across four dimensions to identify which pattern characterizes their situation, enabling pattern-specific strategies. Organizations exhibiting bridging asymmetry should recognize dependence on key individuals and develop institutional mechanisms to reduce vulnerability. Those exhibiting complementary asymmetry face different challenges. The double gap requires developing internal boundary-spanning capability or securing reliable external support. Demonstration over articulation can compensate when value articulation fails. Organizations exhibiting aspirational asymmetry must guard against success-induced complacency, maintaining future-oriented knowledge investment and periodically reassessing gaps between internal tempo and external competitive dynamics.

5.3 Limitations and Future Research

Our study has limitations suggesting future research directions. First, our sample comprises three manufacturing companies implementing enterprise systems. Patterns may not exhaust all configurations, and different patterns may emerge in service industries, public sector organizations, or non-Western contexts. Quantitative research testing the framework across larger samples could validate patterns and identify additional configurations. Second, our longitudinal data spans 15 to 24 months. Longer observation might reveal how patterns evolve. Can organizations transition between patterns as configurations change? Third, our focus on knowledge dimensions may underemphasize other factors such as organizational structure, governance mechanisms, or cultural characteristics. Future research could examine how context moderates relationships between knowledge configurations and alignment outcomes.

6 Conclusions

This study challenges the simplistic view of Shared Domain Knowledge in alignment research. We show that SDK is multidimensional and that knowledge configuration is more important than aggregate level. Our

framework reveals three distinct patterns of knowledge asymmetries that shape alignment processes. Bridging Asymmetry indicates that large gaps can be mitigated through individual dual competencies. Complementary Asymmetry shows that double gaps need external intervention. Aspirational Asymmetry highlights that high shared knowledge can create issues due to expectation-reality gaps and knowledge atrophy. These findings enhance theoretical understanding by identifying mechanisms through which knowledge configurations influence alignment dynamics. For practitioners, the framework provides diagnostic tools for evaluating knowledge configurations and implementing targeted management strategies. Despite limitations, our multidimensional framework offers valuable theoretical insights and practical guidance for managing knowledge asymmetries in business-IT alignment.

7 Ethics declaration

No formal ethical clearance was required for this research. Participants were informed about the study's nature and objectives and provided voluntary consent. Interview data were used for academic research purposes. Company names and individual identities have been anonymized to protect confidentiality. No sensitive or privacy-related information is presented in the study.

8 AI declaration

The present paper was developed by the authors based on the data collected from the case companies. No AI tools were used for developing the paper.

References

Armstrong, C.P. and Sambamurthy, V. (1999) "Information technology assimilation in firms: The influence of senior leadership and IT infrastructures", *Information Systems Research*, Vol. 10, No. 4, pp. 304-327.

Bassellier, G., Benbasat, I. and Reich, B.H. (2001) "Information technology competence of business managers: A definition and research model", *Journal of Management Information Systems*, Vol. 17, No. 4, pp. 159-182.

Brown, C.V. and Grant, G.G. (2005) "Framing the frameworks: A review of IT governance research", *Communications of the Association for Information Systems*, Vol. 15, No. 1, pp. 696-712.

Carlile, P.R. (2004) "Transferring, translating, and transforming: An integrative framework for managing knowledge across boundaries", *Organization Science*, Vol. 15, No. 5, pp. 555-568.

Chan, Y.E. and Reich, B.H. (2007) "IT alignment: what have we learned?", *Journal of Information Technology*, Vol. 22, No. 4, pp. 297-315.

Chun, M. and Mooney, J. (2009) "CIO roles and responsibilities: Twenty-five years of evolution and change", *Information & Management*, Vol. 46, No. 6, pp. 323-334.

Eisenhardt, K.M. (1989) "Building theories from case study research", *Academy of Management Review*, Vol. 14, No. 4, pp. 532-550.

Ewusi-Mensah, K. (1997) "Critical issues in abandoned information systems development projects", *Communications of the ACM*, Vol. 40, No. 9, pp. 74-80.

Gerow, J.E., Grover, V., Thatcher, J.B. and Roth, P.L. (2014) "Looking toward the future of IT-business strategic alignment through the past: A meta-analysis", *MIS Quarterly*, Vol. 38, No. 4, pp. 1059-1085.

Henderson, J.C. and Venkatraman, N. (1993) "Strategic alignment: Leveraging information technology for transforming organizations", *IBM Systems Journal*, Vol. 32, No. 1, pp. 4-16.

Johnson, A.M. and Lederer, A.L. (2010) "CEO/CIO mutual understanding, strategic alignment, and the contribution of IS to the organization", *Information & Management*, Vol. 47, No. 3, pp. 138-149.

Keil, M., Tan, B.C., Wei, K.K., Saarinen, T., Tuunainen, V. and Wassenaar, A. (2000) "A cross-cultural study on escalation of commitment behavior in software projects", *MIS Quarterly*, Vol. 24, No. 2, pp. 299-325.

Langley, A. (1999) "Strategies for theorizing from process data", *Academy of Management Review*, Vol. 24, No. 4, pp. 691-710.

Levina, N. and Vaast, E. (2005) "The emergence of boundary spanning competence in practice: Implications for implementation and use of information systems", *MIS Quarterly*, Vol. 29, No. 2, pp. 335-363.

Luftman, J., Lyytinen, K. and Zvi, T.B. (2017) "Enhancing the measurement of information technology (IT) business alignment and its influence on company performance", *Journal of Information Technology*, Vol. 32, No. 1, pp. 26-46.

Lyytinen, K. and Robey, D. (1999) "Learning failure in information systems development", *Information Systems Journal*, Vol. 9, No. 2, pp. 85-101.

Majchrzak, A., More, P.H. and Faraj, S. (2012) "Transcending knowledge differences in cross-functional teams", *Organization Science*, Vol. 23, No. 4, pp. 951-970.

Miles, M.B. and Huberman, A.M. (1994) *Qualitative Data Analysis: An Expanded Sourcebook*, 2nd ed., Sage Publications, Thousand Oaks, CA.

Orlikowski, W.J. and Yates, J. (2002) "It's about time: Temporal structuring in organizations", *Organization Science*, Vol. 13, No. 6, pp. 684-700.

Pawlowski, S.D. and Robey, D. (2004) "Bridging user organizations: Knowledge brokering and the work of information technology professionals", *MIS Quarterly*, Vol. 28, No. 4, pp. 645-672.

Peppard, J., Edwards, C. and Lambert, R. (2011) "Clarifying the ambiguous role of the CIO", *MIS Quarterly Executive*, Vol. 10, No. 1, pp. 31-44.

Preston, D.S. and Karahanna, E. (2009) "Antecedents of IS strategic alignment: A nomological network", *Information Systems Research*, Vol. 20, No. 2, pp. 159-179.

Reich, B.H. and Benbasat, I. (1996) "Measuring the linkage between business and information technology objectives", *MIS Quarterly*, Vol. 20, No. 1, pp. 55-81.

Reich, B.H. and Benbasat, I. (2000) "Factors that influence the social dimension of alignment between business and information technology objectives", *MIS Quarterly*, Vol. 24, No. 1, pp. 81-113.

Sabherwal, R. and Robey, D. (1995) "Reconciling variance and process strategies for studying information system development", I*nformation Systems Research*, Vol. 6, No. 4, pp. 303-327.

Strauss, A. and Corbin, J. (1998) *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*, 2nd ed., Sage Publications, Thousand Oaks, CA.

Xue, Y., Liang, H. and Boulton, W.R. (2008) "Information technology governance in information technology investment decision processes: The impact of investment characteristics, external environment, and internal context", *MIS Quarterly*, Vol. 32, No. 1, pp. 67-96.

Yin, R.K. (2014) *Case Study Research: Design and Methods*, 5th ed., Sage Publications, Thousand Oaks, CA.

A KM-Based Approach to Service Value Co-Creation and Co-Destruction

Nina Gerges, Paulo Selig, Gregorio Varvakis
EGC/UFSC, Florianopolis, Brazil

Abstract

This study analyzes value co-creation (VCC) and value co-destruction (VCD) in knowledge-intensive services from a knowledge management perspective. According to the dominant service logic, value is co-created in interactions between companies and customers. However, value can also be destroyed by failures in relational, communicative, or resource integration. The study's objective was to propose managerial guidelines to mitigate VCD and promote VCC in this context.
The study employed a qualitative, applied, and prescriptive approach through design science research (DSR). A literature review was conducted, along with seventeen semi-structured interviews with managers, experts, and clients from three small companies operating in the Florianópolis region in the segments of human resources consulting, quality certification, and brand strategy. Based on thematic data analysis, an analytical framework comprising three central dimensions was developed and validated: Communication and Engagement, Knowledge Resources, and Relationship and Collaboration.
The results indicate that communication problems, fragmented channels, low client participation, a lack of knowledge sharing, and difficulty understanding the service act as inhibitors of VCC and catalysts of VCD. Conversely, trust, frequent alignment, clear objectives, access to experts, continuous learning, and collaboration strengthen value creation. Additionally, it was found that team turnover, excessive dependence on experts, and resistance to change increase the risk of value destruction.
As a practical contribution, seven knowledge management (KM)-based management guidelines focused on client leadership engagement, knowledge sharing, organizational learning, and retention of critical knowledge were proposed. In conclusion, KM is essential for mitigating losses and maximizing benefits in knowledge-intensive services.

Keywords: value cocreation, value codestruction; knowledge intensive service, knowledge management; service.

Paper type: Academic Research Paper

1 Introduction

Service organizations, particularly those offering knowledge-intensive services (KIS), have taken on a central role in contemporary economies. In this context, knowledge becomes a strategic asset capable of sustaining competitive differentiation and guiding decisions in complex environments, rather than merely a supporting resource (Grant, 1996; Nonaka and Takeuchi, 1995). Knowledge management (KM) reinforces this perspective, highlighting the creation, integration, and application of knowledge as the foundation of sustainable organizational performance (Nonaka, Krogh, and Voelpel, 2006).

Meanwhile, the dominant service logic has transformed the understanding of value by shifting the focus from unilateral delivery to relational and contextual construction. According to this view, value is not solely produced by the company but is co-created through interactions among those involved in the service. Value is then determined by the beneficiary based on the user experience (Vargo and Lusch, 2016). Customers thus cease to be passive recipients and begin to play an active role in the process.

Co-creation of value (CCV) has been extensively studied for its role in collaboration, engagement, and the integration of competencies between companies and customers. Models such as DART highlight elements such as dialogue, access, transparency, and risk assessment (Prahalad and Ramaswamy, 2004). Other approaches suggest that CCV occurs across multiple touchpoints, depending on reciprocal relationships and the joint mobilization of resources (Payne, Storbacka, and Frow, 2008; Grönroos, 2012).

However, an exclusively positive view is limited because the same process can lead to the co-destruction of value (CDV). Failures in resource integration, knowledge gaps, misaligned expectations, and communication barriers can compromise the service experience and reduce perceived value (Plé and Chumpitaz, 2010). Thus, CDV occurs when interactions cease to generate benefits and begin to produce additional costs for those involved.

Despite theoretical advances, gaps persist in integrating CCV and CDV, particularly in knowledge-intensive services. Therefore, this study proposes a Knowledge Management approach to analyze both phenomena. Through qualitative research grounded in design science research, the study develops an analytical framework

centered on three dimensions: communication and engagement, knowledge resources, and relationships and collaboration. The framework seeks to contribute to theoretical advancement and managerial practice.

2 Literature Review

The literature on services, value, and knowledge management has evolved significantly alongside the increasing significance of services in modern economies. Initially influenced by manufacturing logic, studies began emphasizing quality, customer experience, and the strategic role of services in generating competitive advantage (Heineke and Davis, 2007). Services are now understood as dynamic processes characterized by simultaneity, intangibility, heterogeneity, and perishability, in which customer participation is essential (Fitzsimmons and Fitzsimmons, 2000; Shostack, 1987).

A main paradigm shift occurred with the Dominant Service Logic, which redefined service as applying competencies for another's benefit and positioned value as the result of interactions (Vargo and Lusch, 2007; 2016). From this perspective, customers cease to be passive recipients and begin acting as co-producers, with value determined by their experience (Lusch and Vargo, 2014). Thus, service is understood as an open system in which different actors integrate resources, especially knowledge, to generate benefits (Ballantyne and Varey, 2006).

The concept of value evolves toward the logic of value in use, which is understood as the relationship between the benefits and sacrifices perceived by the customer (Järvi, Kähkönen, and Torvinen, 2018). Value emerges from the interactions between companies and customers and is co-created through the integration of resources and competencies (Vargo, Maglio, and Akaka, 2008; Grönroos, 2012). Value co-creation is a relational, dynamic process in which dialogue, engagement, and knowledge sharing play central roles.

The DART model, comprising dialogue, access, risk assessment, and transparency (Prahalad and Ramaswamy, 2004), stands out among the main theoretical models presented in Figure 1. Payne, Storbacka, and Frow (2008) propose a process-based view that is rooted in company and customer processes and touchpoints. Grönroos (2012), on the other hand, emphasizes that value emerges from the experience of the service itself. All these models acknowledge knowledge as an indispensable strategic resource.

As a counterpoint to the predominantly positive approaches, literature on value co-destruction has emerged. This perspective acknowledges that service interactions can also generate losses. Co-destruction is defined as a process that reduces the well-being of at least one party (Plé and Chumpitaz, 2010) and is associated with failures in resource integration, misaligned expectations, and limitations in participants' capabilities. The main identified factors include a lack of information, a lack of trust, operational errors, inadequate communication, and coordination failures, particularly in B2B relationships (Vafeas, Hughes, and Hilton, 2016). Figure 1 presents the main models and factors related to value co-destruction.

Their consequences include increased costs, loss of customers, and the need to mobilize additional resources to restore service (Plé, 2013). Thus, an integrated review of the literature shows that co-creation and co-destruction are influenced by three central dimensions: communication, knowledge, and relationships. In this context, Knowledge Management plays a strategic role, as the proper creation, sharing, and application of knowledge can enhance value creation and mitigate value destruction.

Elements of Value Co-creation in Services	
Author	Description
Plé and Chumpitaz (2010)	Failure to integrate resources
	Mismatch in expectations
Vafeas, Hughes and Hilton (2016)	Lack of confidence
	Inadequate Communication
	Power imbalance or dependence
Lintula, Tuunanen and Salo (2017)	Objectives and intentions that are divergent or misaligned
	Resources: absence, misuse, or loss
	Mismatch in expectations
Järvi, Kähkönen and Torvinen (2018)	Lack of information
	Trust between the customer and the company
	Lack of clear expectations

Elements of Value Co-Destruction in Services	
Author	Description
Plé and Chumpitaz (2010)	Failure to integrate resources
	Mismatch in expectations
Vafeas, Hughes and Hilton (2016)	Lack of confidence
	Inadequate Communication
	Power imbalance or dependence
Lintula, Tuunanen and Salo (2017)	Objectives and intentions that are divergent or misaligned
	Resources: absence, misuse, or loss
	Mismatch in expectations
Järvi, Kähkönen and Torvinen (2018)	Lack of information
	Trust between the customer and the company
	Lack of clear expectations

Figure 1: Components of the CCV and CDV
Source: prepared by the authors

3 Methodological Procedures

The study aimed to propose guidelines to mitigate co-destruction and promote the co-creation of value in knowledge-intensive services. This is a qualitative, applied, and prescriptive study that seeks to understand organizational phenomena and propose practical solutions (Merriam, 2009; Creswell, 2010).

The methodological approach was based on the Design Science Research (DSR) framework, which is suitable for studies focused on problem-solving through the development of artifacts (Hevner, March, and Park, 2004). Three main artifacts were developed: (i) an analytical framework on the co-creation and co-destruction of value; (ii) a data collection instrument based on semi-structured interviews; and (iii) management guidelines grounded in Knowledge Management.

The methodological process was structured in four stages (Figure 2).

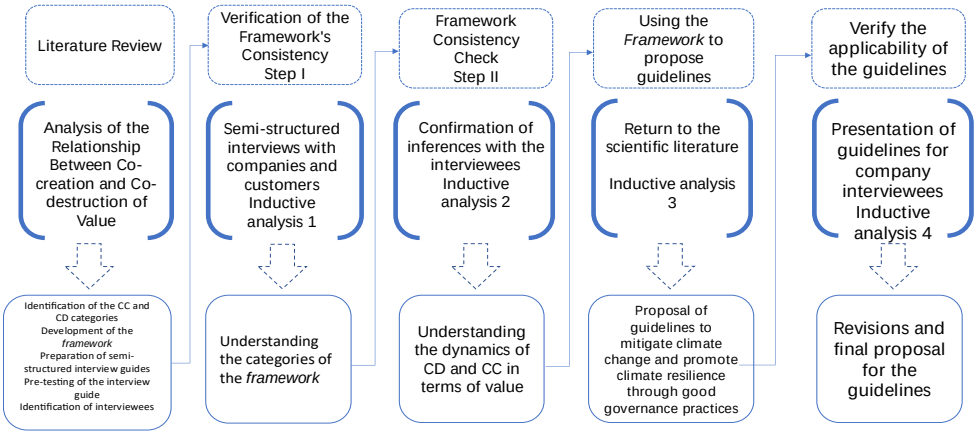


Figure 2: Methodological Process
Source: prepared by the authors

The first phase consisted of a systematic literature review to identify theoretical gaps and related models. Based on inductive analysis, an initial framework comprising three categories was developed: communication and engagement, knowledge resources, and relationships and collaboration. An interview guide was developed and reviewed by experts. Following a pilot test, the guide was refined.

The second stage involved collecting empirical data through 17 interviews conducted between January and July of 2021 with managers, experts, and clients from three small companies in the Florianópolis region that operate in human resources consulting, quality certification, and brand strategy.

The third stage involved analyzing the interview data using thematic analysis and validating the inferences with the participants through triangulation and analytical refinement.

The fourth stage consisted of developing management guidelines by integrating empirical findings with the theoretical framework. Managers from the participating companies evaluated these guidelines, allowing for final adjustments and practical validation.

In summary, the adopted methodological procedures allowed for the articulation of scientific rigor and practical applicability. DSR was appropriate for developing managerial solutions and advancing knowledge about the relationship between knowledge management, co-creation, and co-destruction of value in knowledge-intensive services.

4 Proposed Framework

Gerges (2023) proposed a framework that integrates the theoretical concepts of value co-creation and co-destruction in services. The framework seeks to develop an analytical model to understand and manage interactions between companies and customers (Figure 3). The framework was developed through a systematic review of literature that identified recurring elements in various models, which were then adapted to the context of knowledge-intensive services.

In the field of value co-creation, Prahalad and Ramaswamy (2004) are notable for their emphasis on dialogue, access, risk, and transparency. Payne, Storbacka, and Frow (2008) emphasize service encounters and touchpoints. Grönroos (2012) emphasizes accessibility and communication. Leclercq, Hammedi, and Poncin (2016), on the other hand, incorporate resource integration, engagement, and learning.

Concurrently, studies on the co-creation of value have identified factors that can compromise interaction outcomes. Plé and Chumpitaz (2010) point out failures in resource integration and misaligned expectations. Vafeas, Hughes, and Hilton (2016) point to communication problems, a lack of trust, and power imbalances. Lintula, Tuunanen, and Salo (2017) organize these factors into three dimensions: communication, resources, and expectations. Järvi, Kähkönen, and Torvinen (2018), on the other hand, emphasize informational clarity and relationship quality.

A comparative analysis revealed commonalities between the two fields, particularly regarding communication, resources, knowledge, expectations, trust, and engagement. Based on these findings, the elements were grouped into three categories: (1) communication and engagement, (2) resources and knowledge, and (3) relationships and collaboration.

The first category involves dialogue, information exchange, and conditions that enable interaction. This category includes channels, platforms, and contact environments. Failures in this dimension tend to hinder the co-creation of value.

The second category emphasizes knowledge as the primary resource in knowledge-intensive services. This category encompasses access to knowledge, practical integration by the client, and renewal through learning and innovation generated during interactions.

The third category addresses relational factors that influence the quality of interaction. These factors include alignment of expectations, trust, transparency, collaboration, commitment, and balance of power. Weakened relationships tend to compromise the co-creation of value.

Consolidating these categories resulted in a framework (Figure 3) representing the encounter between company and customer as a dynamic, interdependent space. These categories are complementary and can occur simultaneously throughout the service. The model also recognizes the cumulative nature of interactions, wherein past experiences influence future perceptions.

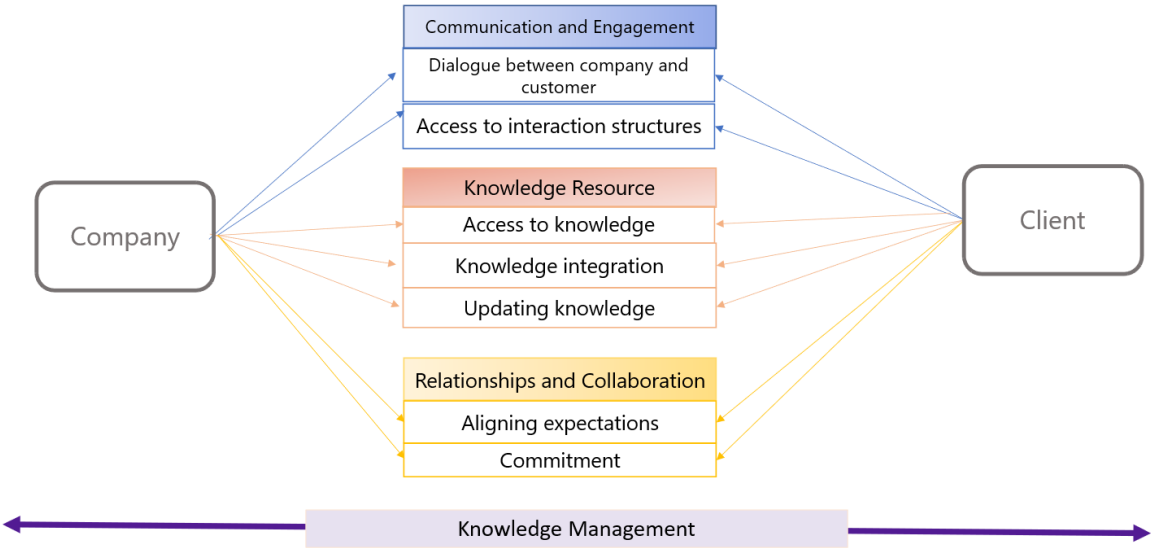


Figure 3: Framework
Source: prepared by the authors

In addition to its analytical function, the framework is prescriptive, providing a basis for guiding questions and management guidelines to mitigate co-destruction and promote the co-creation of value.

5 Discussions

Guiding questions based on the theoretical elements of each category were used to analyze both companies and customers. In the Communication and Engagement dimension, we investigated aspects related to dialogue, the nature of communication between the parties, the main barriers encountered, the channels used, and the ease of accessing contact channels. The Knowledge Resource category aimed to understand how knowledge was shared, accessed, applied, and transformed throughout the service delivery process (Figure 4). The Relationship and Collaboration category addressed trust, alignment of expectations, commitment, transparency, and cooperation among the involved parties. The findings confirm that CCV and CDV emerge from interactions

between customers and companies, as well as from both parties' ability to access, integrate, and utilize resources throughout the service.

The main inhibitors of CCV are concentrated in the Communication and Engagement category. They are particularly related to channel fragmentation, difficulty accessing information, and the multiplicity of decision-makers—factors that increase the complexity of interactions. Furthermore, a lack of customer knowledge, which is linked to the Knowledge Resource category, limits their active participation in the value-creation process. Conversely, the main catalysts for CCV are associated with the Relationship and Collaboration category, including frequent alignment between parties, trust in experts, access to specialists, and clarity of objectives. These factors are reinforced by knowledge sharing, tool use, and learning from service experience.

VDC is mostly caused by issues related to the knowledge resource, such as team turnover, negative past experiences, excessive reliance on experts, and client resistance to change. Difficulties understanding the service and communication failures also contribute to VDC. These results demonstrate that poorly managed knowledge can contribute to value destruction.

Table 1: Data Collection Instrument for the Knowledge Resource Category

Elements	Questions	Authors
Tool for use by the service provider		
Access to knowledge	Do you have any procedures or processes in place to determine whether a customer has the necessary knowledge base for the contracted service?	Prahalad and Ramaswamy (2004); Payne, Storbacka, and Frow (2008); Grönroos (2012)
	Did you help the client gain access to the basic knowledge required for the service? How?	
Knowledge integration	How did you help the client apply the knowledge provided through the service within their company?	Plé and Chumpitaz, 2010; Lintula, Tuunanen, and Salo, 2017
	What obstacles did the client face in applying the knowledge provided through the service?	
Updating knowledge	Do you know if the client gained new insights from the service provided?	Leclercq, Hammedi, and Poncin (2016)
	In your opinion, what obstacles prevented the client from gaining new insights? How did you help them with that?	
Tool for use by the client		
Access to knowledge	Did your company receive any training or background information before hiring the service?	Prahalad and Ramaswamy (2004); Payne, Storbacka, and Frow (2008); Grönroos (2012)
	Did Company X provide you with the necessary expertise to perform the service? How?	
Knowledge integration	Has your company been able to apply the knowledge provided through the service?	Plé and Chumpitaz, 2010; Lintula, Tuunanen, and Salo, 2017
	How did you go about applying that knowledge in your company? Did Company X help? How?	
Updating knowledge	Do you think your company has gained new insights from the work performed? Can you give an example?	Leclercq, Hammedi, and Poncin (2016)
	Did Company X assist with this process? How did that happen?	

Source: prepared by the authors

The three categories of the framework operate simultaneously and interdependently, forming a dynamic system of interactions. In this environment, episodes of CCV and CDV can coexist and alternate throughout the service, necessitating continuous, adaptive management. Seven guidelines based on knowledge management emerge from this systemic understanding, as shown in Figure 4.

The first guideline involves engaging the client's leadership, reducing resistance, and aligning expectations by actively involving decision-makers. The second guideline is to ensure expert participation to facilitate access to knowledge and reduce critical dependencies. The third guideline proposes fostering the use of information and communication systems and structuring information flows and shared repositories. The fourth guideline involves knowledge sharing to reduce asymmetries between the company and the client and promote collaboration. The fifth guideline seeks to stimulate continuous learning throughout the service to strengthen competencies and innovation. The sixth guideline involves continually iterating on customer needs and service objectives to enable adjustments in response to contextual changes. The seventh guideline recommends retaining and integrating team knowledge to minimize the impact of turnover and preserve organizational memory.

Figure 4: Guidelines for Mitigating CDV and Promoting CCV in Knowledge-Intensive Services
Source: prepared by the authors

Taken together, the guidelines translate empirical findings into practical guidance and demonstrate that the effective management of knowledge-intensive services depends on the interplay between communication, knowledge, and relationships. Rather than isolated actions, these are complementary mechanisms that transform complex interactions into ongoing opportunities to co-create value and mitigate co-destruction.

6 Final Considerations

This study analyzed value co-creation (VCC) and value co-destruction (VCD) in knowledge-intensive services from a knowledge management perspective. It validated a three-category framework: Communication and Engagement, Knowledge Resources, and Relationships and Collaboration. The results demonstrate that VCC and VCD arise from the interactions between companies and customers and are influenced by how knowledge is accessed, shared, and applied throughout the service.

From a theoretical standpoint, failures in communication and engagement were found to inhibit VCC, while the absence of knowledge sharing was found to be the main catalyst for VCD. Conversely, strengthening relationships and collaboration, particularly through trust and alignment, fosters value creation. The study also highlights the roles of client leadership and experts, as well as challenges related to information systems, learning, and turnover.

As a key contribution, seven guidelines based on GC were proposed: engage client leadership, ensure expert participation, foster information and communication systems, share knowledge, encourage learning, iterate service needs and objectives, and retain team knowledge.

In practice, these guidelines allow for the systematization of existing actions and the incorporation of new management practices. A limitation of this study is its empirical scope. Future studies may broaden the scope of contexts and utilize quantitative and technological approaches. In conclusion, GC is essential for transforming interactions into continuous processes of shared value generation.

7 Acknowledgements

This study was supported by the CNPq Brazilian Government agency and NGS/UFSC by FEPESE funding

8 Ethics declaration

Considering the requirements of our university's ethics committee, the participants were informed and attested to their free participation. Therefore, the work is free for publication.

9 AI declaration

AI/GPT was used to format the references as requested by IFKAD.

References

Ballantyne, D. and Varey, R.J. (2006) 'Creating value-in-use through marketing interaction: the exchange logic of relating, communicating and knowing', *Marketing Theory*, Vol. 6, No. 3, pp. 335–348.

Creswell, J.W. (2010) *Projeto de pesquisa: métodos qualitativos, quantitativos e mistos*, 3rd ed., Porto Alegre: Artmed.

Gerges, N.R.C. (2023). *Uma proposta para mitigar ou promover a codestruição de valor do serviço por meio da gestão do conhecimento.* Tese de doutorado, Universidade Federal de Santa Catarina.

Grant, R.M. (1996) 'Toward a knowledge-based theory of the firm', *Strategic Management Journal*, Vol. 17, No. 2, pp. 109–122.

Gronroos, C. (1990) 'Relationship approach to marketing in service contexts: the marketing and organizational behavior interface', *Journal of Business Research*, Vol. 20, No. 1, pp. 3–11.

Grönroos, C. (2012) 'Conceptualising value co-creation: a journey to the 1970s and back to the future', *Journal of Marketing Management*, Vol. 28, No. 13–14, pp. 1520–1534.

Heineke, J. and Davis, M.M. (2007) 'The emergence of service operations management as an academic discipline', *Journal of Operations Management*, Vol. 25, pp. 364–374.

Hevner, A.R., March, S.T. and Park, J. (2004) 'Design science in information systems research', *MIS Quarterly*, Vol. 28, No. 1, pp. 75–105.

Järvi, H., Kähkönen, A.K. and Torvinen, H. (2018) 'When value co-creation fails: Reasons that lead to value co-destruction', *Scandinavian Journal of Management*, Vol. 34, No. 1, pp. 63–77.

Leclercq, T., Hammedi, W. and Poncin, I. (2016) 'Ten years of value cocreation: an integrative review', *Recherche et Applications en Marketing*, Vol. 31, No. 3, pp. 26–60.

Lintula, J., Tuunanen, T. and Salo, M. (2017) 'Conceptualizing the value co-destruction process for service systems: literature review and synthesis', in *Proceedings of the 50th Hawaii International Conference on System Sciences (HICSS)*, pp. 1632–1641.

Lusch, R.F. and Vargo, S.L. (2014) *Service-Dominant Logic: premises, perspectives, possibilities*, New York: Cambridge.

Lusch, R.F. and Webster, F.E. (2011) 'A stakeholder-unifying, cocreation philosophy for marketing', *Journal of Macromarketing*, Vol. 31, No. 2, pp. 129–134.

Merriam, S.B. (2009) Qualitative research: a guide to design and implementation. San Francisco: Jossey-Bass.

Nonaka, I., Krogh, G.V. and Voelpel, S. (2006) 'Organizational knowledge creation theory: evolutionary paths and future advances', *Organization Studies*, Vol. 27, No. 8, pp. 1179–1208.

Nonaka, I. and Takeuchi, H. (1995) *The knowledge-creating company: how Japanese companies create the dynamics of innovation*, Oxford: Oxford University Press.

Payne, A.F., Storbacka, K. and Frow, P. (2008) 'Managing the co-creation of value', *Journal of the Academy of Marketing Science*, Vol. 36, No. 1, pp. 83–96.

Plé, L. (2013) 'How does the customer fit in relational coordination? An empirical study in multichannel retail banking', *Management*, Vol. 16, No. 1, pp. 1–30.

Plé, L. (2016) 'Studying customers resource integration by service employees in interactional value co-creation', *Journal of Services Marketing*, Vol. 30, No. 2, pp. 152–164.

Plé, L. (2017) 'Why do we need research on value co-destruction?', *Journal of Creating Value*, Vol. 3, No. 2, pp. 162–169.

Plé, L. and Chumpitaz, C.R. (2010) 'Not always co-creation: introducing interactional co-destruction of value in service-dominant logic', *Journal of Services Marketing*, Vol. 24, No. 6, pp. 430–437.

Plé, L. and Lecocq, X. (2015) 'Customers as creative resources: their influence on firm freedom', *Journal of Business Strategy*, Vol. 36, No. 4, pp. 1–22.

Prahalad, C.K. and Ramaswamy, V. (2000) 'Co-opting customer competence', *Harvard Business Review*, Vol. 78, pp. 79–90.

Prahalad, C.K. and Ramaswamy, V. (2004) 'Co-creating unique value with customers', *Strategy and Leadership*, Vol. 32, No. 3, pp. 4–9.

Shostack, G.L. (1987) 'Service positioning through structural change', *The Journal of Marketing*, Vol. 51, pp. 34–43.

Vafeas, M., Hughes, T. and Hilton, T. (2016) 'Antecedents to value diminution: a dyadic perspective', *Marketing Theory*, Vol. 16, No. 4, pp. 469–491.

Vargo, S.L. and Lusch, R.F. (2007) 'Why service?', *Journal of the Academy of Marketing Science*, Vol. 36, No. 1, pp. 25–38.

Vargo, S.L. and Lusch, R.F. (2016) 'Institutions and axioms: an extension and update of service-dominant logic', *Journal of the Academy of Marketing Science*, Vol. 44, pp. 5–23.

Vargo, S.L., Maglio, R.P. and Akaka, M.A. (2008) 'On value and value co-creation: a service systems and logic perspective', *European Management Journal*, Vol. 26, pp. 145–152.

Artificial Intelligence and Marketing Creativity: A Thematic Synthesis of Disruption, Human-AI Collaboration, and Strategic Transformation

Olufemi Afolabi[1], Oluseye Jegede[2]
[1] Global Banking School, Leeds, United Kingdom
[2] Footprint School of Business, London, United Kingdom

Abstract

This study examines the transformative role of artificial intelligence in reshaping marketing creativity, with particular emphasis on human–AI collaboration, organisational capability, and strategic outcomes. Drawing on a theory-driven thematic synthesis of thirteen peer-reviewed journal articles and academic theses, the paper advances a socio-technical perspective in which creativity is no longer viewed as an exclusively human attribute but as an emergent outcome of interaction between intelligent systems, human judgement, and institutional contexts. The analysis integrates Innovation Diffusion Theory, the Technology Acceptance Model, and socio-technical systems theory to examine how AI capabilities, human creative agency, organisational design, and governance mechanisms jointly produce marketing creativity outcomes.
The findings indicate that AI significantly expands the scope, speed, and diversity of idea generation, thereby redefining the early stages of the creative process. Generative AI systems enable rapid production of alternative headlines, visuals, layouts, and narratives, fundamentally altering ideation capacity and lowering the marginal cost of creative experimentation. However, increased ideation does not necessarily translate into enhanced creative value. Human agency remains central, mediating the relationship between AI capability and meaningful marketing outcomes through evaluative judgement, narrative construction, and strategic alignment. Without deliberate human involvement, AI-generated outputs tend toward statistical normativity, limiting originality and emotional depth. Human evaluative competence thus becomes the critical bottleneck in converting generative abundance into brand-relevant, culturally sensitive, and strategically coherent creative outputs.
The study demonstrates that organisational readiness, including AI literacy and evaluative competence, critically shapes the effectiveness of AI-augmented creativity. Firms with stronger strategic and evaluative skills are better positioned to translate AI-generated ideas into coherent campaigns, avoiding homogenisation and ethical missteps. Ethical governance and legitimacy emerge as foundational conditions influencing both market acceptance and performance. Issues of privacy, transparency, authenticity, and authorship actively shape consumer trust and the symbolic legitimacy of brand communications, forming part of the creative system rather than external constraints.
The paper argues that AI does not diminish creativity but redistributes it, shifting competitive advantage from idea generation to interpretation, coordination, and ethical stewardship. Under AI augmentation, creativity increasingly resides in the selection, shaping, legitimisation, and strategic deployment of ideas rather than in their initial production. The study concludes that innovation-driven, human-centred integration of AI is essential for sustained differentiation and strategic transformation. Organisations should prioritise evaluative capability development, design human–AI collaboration with clear role delineation, and embed ethical governance within creative workflows. In this reconfigured landscape, creativity remains central to marketing success, residing in human judgement, organisational design, and ethical stewardship within AI-enabled marketing systems.

Keywords: Artificial Intelligence, Marketing Creativity, Human–AI Collaboration, Global Innovation Systems, Organisational Capability, Ethical Governance

Paper Type: Academic Research Paper

1 Introduction

Creativity has long been positioned as a core human capability underpinning marketing and advertising effectiveness. Traditionally, creative advantage has been associated with imagination, intuition, cultural sensitivity, and narrative judgement, supported by qualitative insight and experiential knowledge. However, the rapid diffusion of artificial intelligence (AI), particularly generative and predictive systems has fundamentally altered how creative work is produced, evaluated, and deployed within marketing systems (Pagani & Wind, 2024; Davenport et al., 2023).

Historical waves of technological disruption in advertising from print and broadcast media to digital platforms and programmatic advertising have consistently reshaped creative production and media strategy. AI represents the latest and arguably most transformative phase in this trajectory, enabling the generation of copy, visuals, and video content at unprecedented speed and scale (Huang & Rust, 2021). While early marketing

technology research framed AI primarily as an efficiency-enhancing tool for segmentation, targeting, and optimisation (McDowall, 2018), more recent scholarship suggests a deeper creative reconfiguration, with AI increasingly involved in ideation, experimentation, and content production (Pagani & Wind, 2024).

At the strategic level, AI is widely characterised as a disruptive technology capable of reshaping organisational structures, creative workflows, and competitive dynamics (Păvăloaia & Necula, 2023). Global expenditure on AI-enabled advertising tools is projected to exceed US$100 billion by 2030 (Statista, 2024), reflecting widespread industry adoption. However, despite this enthusiasm, empirical clarity regarding AI's impact on *creative quality*—as opposed to operational efficiency—remains limited. Industry metrics often conflate automation speed with genuine creative effectiveness, obscuring questions of originality, emotional resonance, and authenticity.

Crucially, the integration of AI into advertising creativity is not frictionless. Conventional creative practice relies heavily on human perception, intuition, and narrative sensemaking, which do not naturally align with data-driven, algorithmic systems (Lee & Cho, 2020; Zaki, 2019). Over-reliance on automation risks producing predictable and emotionally shallow outputs. Yet, when applied in partnership with human judgement, AI can reveal latent insights, suggest novel combinations, and enhance creative productivity (Rust, 2020). Understanding how AI can be integrated creatively—rather than merely efficiently—therefore represents a central theoretical and practical challenge for contemporary marketing.

Despite a growing body of work, existing research remains fragmented across marketing, information systems, psychology, design, and futures studies. Empirical findings are heterogeneous, and theoretical integration is limited. This study addresses these gaps through a theory-driven thematic synthesis of thirteen peer-reviewed articles and academic theses at specifically explore the context of artificial intelligence and marketing creativity, guided by the following research questions:

- *How is AI reshaping creativity in marketing and advertising?*
- *What forms of human–AI collaboration emerge across creative contexts?*
- *What strategic, organisational, and ethical implications arise from AI-driven creative disruption?*

2 Literature Review

2.1 AI as a disruptive force in marketing and advertising

AI is widely recognised as a general-purpose disruptive technology whose implications extend beyond efficiency gains to strategic transformation (Păvăloaia & Necula, 2023). In advertising, disruption manifests through automated content production, predictive analytics, and real-time personalisation at a scale unattainable through traditional methods (Huang & Rust, 2021).

From a diffusion perspective, the adoption of AI reflects patterns described by Innovation Diffusion Theory (IDT), including perceived relative advantage, compatibility with existing creative workflows, complexity, trialability, and observability (Rogers, 1962). While AI offers clear operational advantages, adoption rates vary considerably across organisations due to concerns over creative autonomy, loss of originality, and cultural fit within creative teams (Huang & Rust, 2021). Trialability—such as piloting AI tools on limited campaigns—has emerged as a key mechanism for reducing perceived risk and enabling gradual integration.

Complementing this organisational view, the Technology Acceptance Model (TAM) highlights the importance of individual perceptions of usefulness and ease of use in shaping AI adoption (Davis, 1989). In creative advertising contexts, perceived usefulness extends beyond productivity to include perceived contribution to idea quality, relevance, and engagement. Perceived ease of use is particularly salient, as complex tools can generate resistance among creative professionals, undermining adoption even where strategic benefits are evident (Stone et al., 2020).

Integrating IDT and TAM provides a more holistic understanding of AI adoption in advertising, capturing both organisational diffusion dynamics and individual creative acceptance. However, neither framework alone adequately addresses questions of creative agency, authorship, and emotional resonance, necessitating a broader socio-technical perspective.

2.2 Creativity, AI, and advertising practice

Creativity in marketing is traditionally defined by outputs that are both novel and useful. AI complicates this definition by introducing non-human agents into the creative process. Rather than replacing creativity, AI increasingly functions as a collaborator—supporting ideation, variation, and testing, while humans retain responsibility for evaluation, narrative coherence, and strategic alignment (Rust, 2020; Pagani & Wind, 2024).

Empirical studies in advertising contexts show that generative AI can alleviate cognitive constraints such as "blank-page" anxiety and accelerate experimentation (Albakry et al., 2025). Predictive analytics and recommendation systems further support creative decision-making by identifying patterns in consumer behaviour and preferences (Pradeep et al., 2018). However, when creative production becomes overly dependent on algorithmic outputs, results may converge towards statistically average forms, reducing distinctiveness and emotional depth (Rathore, 2016).

These findings reinforce the view of creativity as a socio-technical phenomenon, emerging from interaction between humans, algorithms, organisational routines, and governance structures (Ameen et al., 2022). Effective creative integration therefore depends not on technological sophistication alone, but on how human–AI collaboration is designed and managed.

2.3 Human–AI collaboration and creative agency

Across the reviewed literature, a recurring theme is the reconfiguration of creative agency. Pagani and Wind (2024) highlight tensions between autonomy and assistance, as creators negotiate control over AI-generated outputs. Similarly, Ameen et al. (2022) caution that creativity may become constrained if AI-generated "defaults" dominate decision-making.

Amankwah-Amoah et al. (n.d.) extend this concern to the creative industries more broadly, arguing that authenticity and the "human touch" may become strategic differentiators in AI-saturated markets. These findings suggest that collaboration design—how tasks and authority are allocated between humans and AI—is central to creative outcomes.

2.4 Governance, ethics, and legitimacy

Ethical concerns cut across all strands of the literature. McDowall (2018) emphasises privacy and data governance as existential issues for future marketing systems. Albakry et al. (2025) highlight originality and academic integrity concerns in AI-supported creative education, while Amankwah-Amoah et al. (n.d.) stress legitimacy and trust in creative outputs.

Despite consensus on the importance of governance, the literature offers limited operational guidance, revealing a critical gap between normative concerns and practical implementation.

3 Theoretical Framework and Propositions

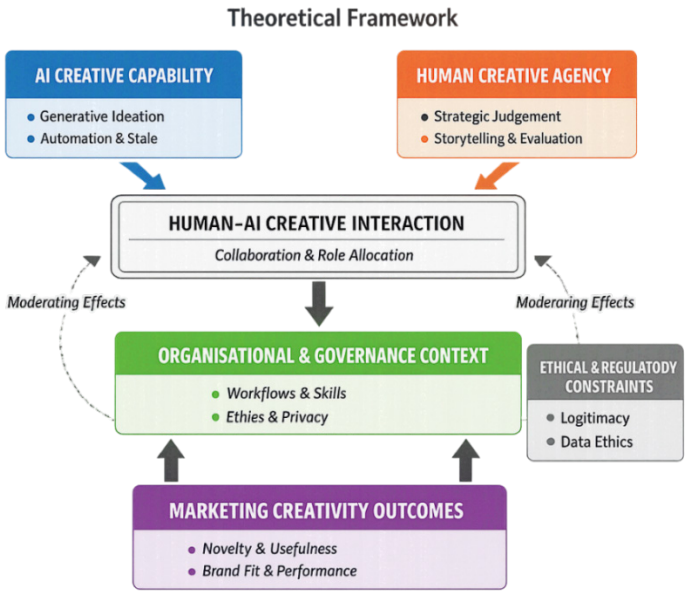
3.1 Conceptual foundations

Drawing on marketing creativity theory, disruptive innovation theory, and socio-technical systems theory, this study conceptualises AI-augmented marketing creativity as an *emergent outcome of human–AI collaboration embedded within organisational and governance structures*. Rather than treating AI as a standalone technological input, the framework positions creativity as jointly produced through interactions among AI capabilities, human judgement, organisational design, and ethical–regulatory constraints (Ameen et al., 2022; Pagani & Wind, 2024; McDowall, 2018).

The framework integrates four core constructs:

1. **AI Creative Capability** (e.g. generative ideation, variation, recombination, speed, scale)
2. **Human Creative Agency** (e.g. strategic framing, evaluative judgement, narrative sensemaking)
3. **Organisational & Governance Context** (e.g. workflows, skills, ethics, privacy, legitimacy)
4. **Marketing Creativity Outcomes** (e.g. novelty, usefulness, brand fit, market performance)

3.2 Theoretical framework diagram (textual representation)



Source: Authors

3.3 Propositions Development

Proposition 1 (P1): AI Creative Capability and Ideation Expansion

AI creative capability positively influences the volume and diversity of marketing ideas generated, thereby expanding the ideation space available to marketers.

Rationale:

Generative AI excels at rapid variation and recombination of creative elements, enhancing divergent thinking in advertising contexts (Pagani & Wind, 2024; Pradeep et al., 2018). However, novelty at scale does not equate to creative value without human evaluation.

Proposition 2 (P2): Human Creative Agency as Mediator

Human creative agency mediates the relationship between AI creative capability and marketing creativity outcomes.

Rationale:

AI lacks contextual sensitivity to brand meaning, cultural nuance, and emotional storytelling. Human judgement remains essential for framing problems, assessing relevance, and legitimising creative outputs (Ameen et al., 2022; Rust, 2020).

Proposition 3 (P3): Collaboration Design and Role Clarity

Clearly defined human–AI collaboration roles strengthen the positive effect of AI creative capability on marketing creativity outcomes.

Rationale:

Advertising studies show that creativity improves when AI is positioned as an ideation assistant rather than a decision authority, preserving human autonomy and evaluative control (Pagani & Wind, 2024; Albakry et al., 2025).

Proposition 4 (P4): Organisational Capability as Moderator

Organisational AI literacy and evaluative capability positively moderate the relationship between human creative agency and marketing creativity outcomes.

Rationale:

Firms with stronger strategic and evaluative skills are better able to translate AI-generated ideas into coherent campaigns, avoiding homogenisation and ethical missteps (AbouElgheit, 2025; McDowall, 2018).

Proposition 5 (P5): Governance and Legitimacy

Ethical governance mechanisms moderate the impact of AI-augmented creativity on market performance.

Rationale:

Consumer trust, perceptions of authenticity, and compliance with privacy and transparency norms shape acceptance of AI-enabled advertising (Amankwah-Amoah et al., n.d.; McDowall, 2018).

Proposition 6 (P6): Disruptive Competitive Outcomes

AI-augmented marketing creativity enables disruptive business model innovation when integrated with strategic repositioning.

Rationale:

AI lowers creative production costs and entry barriers, benefiting smaller firms, but sustained advantage depends on complementary organisational change (AbouElgheit, 2025; Păvăloaia & Necula, 2023).

Integrated Contribution

By integrating advertising-specific adoption theories (IDT, TAM) with socio-technical creativity frameworks, this study advances marketing theory in three ways:

- **Conceptual:** Reframes advertising creativity as a human–AI system outcome rather than an individual capability.
- **Strategic:** Explains how AI-enabled creativity contributes to disruption without eroding differentiation.
- **Governance-oriented:** Positions ethics and legitimacy as central moderators of creative value.

4 Methodology

4.1 Research design

This study adopts a qualitative, theory-driven thematic analysis to examine how artificial intelligence reshapes creativity within marketing and advertising contexts. The methodological approach is explicitly guided by the proposed theoretical framework, which conceptualises marketing creativity as an outcome of human–AI interaction, shaped by organisational capabilities and governance conditions.

Given the exploratory and integrative nature of the research questions, a qualitative synthesis was considered more appropriate than a quantitative meta-analysis. The reviewed studies vary considerably in design, scope, and empirical focus, and do not report comparable effect sizes. Instead, this study seeks to identify patterns of meaning, convergence, and tension across the literature, consistent with interpretive approaches commonly used in marketing theory development.

4.2 Data sources and selection criteria

The analysis draws on thirteen academic sources that specifically looked at the interphase between artificial intelligence and marketing creativity, including peer-reviewed journal articles, systematic reviews, and examined theses. These sources were selected using three criteria:

- Explicit engagement with artificial intelligence in marketing, advertising, or creative industries
- Conceptual or empirical treatment of creativity, innovation, or strategic transformation
- Scholarly credibility and relevance to contemporary marketing practice

The inclusion of doctoral theses alongside journal articles enabled the incorporation of future-oriented and practice-embedded insights that are often under-represented in journal literature.

4.3 Analytical Procedure

The analysis followed a structured yet flexible six-stage thematic process, informed by the theoretical framework:

- Familiarisation through repeated reading of each source
- Initial coding of key concepts linked to AI capability, human agency, creativity outcomes, organisational context, and governance
- Axial coding, linking concepts to the relationships proposed in the framework
- Theme refinement, ensuring consistency across sources
- Framework mapping, aligning themes with propositions P1–P6
- Interpretive synthesis, integrating findings into a coherent explanatory narrative

Throughout the process, attention was paid not only to areas of agreement but also to contradictions and boundary conditions, strengthening analytical depth.

5 Results and Discussion

This section presents and discusses the findings of the thematic analysis in direct relation to the theoretical framework. Rather than treating results and discussion as separate stages, the analysis is integrated, allowing empirical insights from the reviewed literature to be interpreted through the lens of the proposed propositions

(P1–P6). Academic sources are embedded within each subsection to demonstrate how existing scholarship informs, supports, or nuances the findings.

5.1 AI Creative capability and the expansion of ideation (P1)

Across the reviewed literature, there is strong agreement that AI significantly expands the volume, speed, and diversity of creative ideas available to marketers. Generative AI systems enable rapid production of alternative headlines, visuals, layouts, and narratives, fundamentally altering the early stages of the creative process. Pagani and Wind (2024) show that AI tools increase exploratory breadth in marketing creativity by enabling users to test multiple creative directions in parallel, something that would be impractical using purely human labour.

Similarly, Albakry et al. (2025) demonstrate that in advertising and digital content design contexts, AI-supported ideation enhances creative confidence and lowers psychological barriers associated with "blank-page" problems. These findings align with broader observations in the disruptive technology literature, which positions AI as a catalyst for efficiency and experimentation across business functions (Păvăloaia & Necula, 2023).

However, the literature also reveals a consistent qualification: greater ideation capacity does not guarantee superior creativity. While AI can generate novel combinations at scale, it lacks contextual sensitivity to brand values, cultural nuance, and strategic intent. As Ameen et al. (2022) argue, novelty alone is insufficient in marketing contexts where usefulness and relevance are equally central to creative effectiveness. The findings therefore support Proposition 1 while reinforcing the need to distinguish between idea quantity and creative value.

5.2 Human creative agency as a mediating mechanism (P2)

The analysis highlights human creative agency as a central mediating force between AI capability and marketing creativity outcomes. Across conceptual, empirical, and futures-oriented studies, humans retain responsibility for problem framing, evaluative judgement, and narrative sensemaking, even as AI contributes to idea generation.

Ameen et al. (2022) emphasise that creativity in AI-augmented environments emerges from interaction rather than substitution. Their theoretical work argues that without deliberate human involvement, AI-generated outputs tend towards statistical normativity, limiting originality in strategic contexts. Pagani and Wind (2024) provide empirical support for this claim, showing that marketers who actively engaged in iterative evaluation and refinement produced more distinctive and strategically aligned creative outputs than those who relied heavily on default AI suggestions.

The reviewed theses further reinforce this point. McDowall (2018) anticipates that while automation will reshape marketing work, human judgement will remain central in navigating complexity, ambiguity, and ethical trade-offs. Collectively, these findings strongly support Proposition 2, positioning human agency as an indispensable mechanism through which AI-enabled creativity is translated into meaningful marketing outcomes.

5.3 Human–AI collaboration design and role clarity (P3)

Beyond individual agency, the findings suggest that how collaboration between humans and AI is structured plays a critical role in shaping creative effectiveness. Studies consistently indicate that creativity is enhanced when roles are clearly delineated, with AI positioned as a support for ideation and variation, and humans retaining control over selection and strategic direction.

Pagani and Wind (2024) observe that ambiguity in role allocation can lead to over-reliance on AI-generated outputs, reducing critical engagement and creative confidence. This is echoed in educational research by Albakry et al. (2025), where students produced stronger creative work when AI was used as a co-creative assistant rather than an authoritative source.

These findings align with socio-technical perspectives on creativity, which emphasise the importance of system design in shaping outcomes (Ameen et al., 2022). Proposition 3 is therefore supported: clear collaboration design strengthens the positive relationship between AI capability and creative outcomes, while poorly designed interaction risks undermining creative quality.

5.4 Organisational capabilities and skill readiness (P4)

The analysis further reveals that organisational context significantly conditions the effectiveness of AI-augmented creativity. Firms with higher levels of AI literacy, strategic competence, and evaluative capability were better positioned to harness AI-generated ideas productively.

AbouElgheit (2025) demonstrates that small marketing firms can leverage generative AI to disrupt established competitors, but only when AI adoption is combined with strategic repositioning and skill development. This finding resonates with McDowall's (2018) argument that technological innovation alone is insufficient; organisational learning and capability building are essential for sustained advantage.

The literature also suggests that evaluative skills—such as assessing creative fit, ethical risk, and market relevance—become more important as AI lowers the cost of idea generation. These findings provide strong support for Proposition 4, highlighting organisational capability as a positive moderator of human creative agency.

5.5 Governance, ethics, and market legitimacy (P5)

Ethical and governance considerations emerge as structural conditions, rather than secondary concerns, shaping the outcomes of AI-enabled creativity. Issues related to data privacy, authenticity, transparency, and authorship recur across the literature.

McDowall (2018) identifies privacy regulation and public trust as key uncertainties influencing the future of marketing innovation. In creative contexts, Amankwah-Amoah et al. (n.d.) argue that excessive automation risks eroding authenticity, potentially undermining consumer trust in creative industries. Albakry et al. (2025) similarly highlight concerns around originality and ethical use of AI-generated content in educational and creative practice.

These insights collectively support Proposition 5, suggesting that governance mechanisms moderate the relationship between creativity and market performance. Where ethical oversight is embedded within creative workflows, AI-enhanced creativity is more likely to be perceived as legitimate and trustworthy.

5.6 AI-Augmented creativity and disruptive outcomes (P6)

The findings indicate that AI-enabled creativity can contribute to disruptive innovation, particularly for smaller and more agile organisations. By reducing creative production costs and accelerating experimentation, AI opens opportunities for new market entrants and alternative service models (AbouElgheit, 2025).

However, disruption is shown to be contingent on strategic integration. Firms that adopted AI tools without revising their business models or value propositions achieved limited competitive benefit. This finding is consistent with disruptive innovation theory and with broader assessments of AI as a general-purpose technology whose impact depends on complementary organisational change (Păvăloaia & Necula, 2023).

Proposition 6 is therefore supported, with the important caveat that AI-enabled creativity only becomes disruptive when embedded within coherent strategic and organisational transformation.

6 Integrated Discussion and Theoretical Contribution

Taken together, the findings indicate that AI does not erode the importance of creativity in marketing; rather, it redistributes creative value across the marketing system. Where creativity was previously treated as a predominantly human capability expressed through ideation and artistic execution, it is increasingly constituted through socio-technical interaction, a coordinated assemblage of generative tools, human evaluative judgement, organisational routines, and governance mechanisms. In this sense, AI shifts the creative "centre of gravity" from the production of ideas to the selection, shaping, legitimisation, and strategic deployment of ideas, thereby reconfiguring the locus of advantage from *creative generation* to *creative governance and sensemaking* (Ameen et al., 2022; Pagani & Wind, 2024).

6.1 Reconfiguring the creative production function: from scarcity to abundance

A central implication concerns the transformation of the creative production function under conditions of idea abundance. Generative models reduce the marginal cost of producing copy, layouts, imagery, and variations close to zero, enabling continuous iteration and large-scale experimentation. This creates a structural inversion: historically, ideation was scarce and evaluation comparatively manageable; with AI, ideation becomes plentiful while evaluation becomes the bottleneck. This bottleneck is not simply managerial, it is epistemic. Human actors

must determine what is *good*, *appropriate*, *distinctive*, and *brand-consistent* in environments where the system can output thousands of plausible artefacts quickly.

This shift helps explain why many organisations report gains in efficiency without commensurate gains in distinctiveness. Generative systems are trained to reproduce statistically likely patterns, which can elevate baseline competence while increasing the risk of normative convergence, creative outputs that are fluent and polished yet strategically generic. The implication is that competitive advantage will increasingly depend on differentiated evaluative capacity: the ability to impose a distinctive interpretive frame, curate outputs with cultural and brand sensitivity, and translate ideas into coherent narratives rather than simply generating content at scale.

6.2 Human creative agency as strategic control, not residual involvement

The results support the view that human agency is not a symbolic residue left after automation; it is the principal mechanism of strategic control in AI-augmented creativity. Specifically, humans retain responsibility for problem framing (deciding what the campaign is *for*), evaluative judgement (deciding what *counts as quality*), and narrative sensemaking (deciding what the idea *means* for the brand and its audiences). These functions are difficult to formalise because they involve tacit knowledge, cultural interpretation, and value-based decision-making. In marketing, such judgements are inseparable from questions of legitimacy: what a brand can credibly say, what an audience will find authentic, and what constitutes acceptable persuasion.

Consequently, the contribution of AI to creativity is conditional. AI can expand the ideation space and accelerate iteration, but its outputs require human governance to avoid hollow novelty (new but irrelevant) or fluent sameness (well-written but interchangeable). This aligns with socio-technical perspectives that treat creativity as emergent from interaction rather than located in a single agent (Ameen et al., 2022). It also clarifies why creativity increasingly resides in coordination: the orchestration of tools, people, and standards across the creative pipeline.

6.3 Collaboration design as a micro-foundation of creative performance

A deeper theoretical implication is that collaboration design, how roles, authority, and accountability are allocated between humans and AI functions as a micro-foundation of marketing creativity outcomes. The findings suggest that "AI as generator, human as evaluator" is not merely a convenient division of labour; it constitutes a mechanism that protects autonomy, sustains critical engagement, and mitigates default dependency. When role boundaries are ambiguous, creative teams may accept AI outputs as authoritative suggestions, reducing critical scrutiny and narrowing exploration. In contrast, when AI is structured as an exploratory engine and humans as interpretive governors, teams are more likely to use AI outputs as prompts for divergence rather than templates for replication.

This insight extends system-level marketing creativity theory by specifying *how* socio-technical systems shape outcomes: not through AI adoption per se, but through interaction protocols (briefing practices, prompt governance, review routines, approval rights, and feedback loops). These organisational "interaction grammars" determine whether AI increases variety and distinctiveness or merely amplifies speed and volume.

6.4 Organisational capability: the rise of evaluative competence and creative ops

At the organisational level, the findings imply that AI shifts marketing capability requirements away from pure creative talent and towards evaluative competence, AI literacy, and what might be termed *creative operations* (creative ops): the routines and infrastructures that manage high-throughput creative experimentation. Under AI augmentation, firms must develop capacity for:

- Brand constraint articulation (codifying tone, values, exclusions, and boundaries in usable form)
- Quality assurance systems (human review, bias checks, factual verification, and compliance screening)
- Performance learning loops (linking creative variants to engagement, conversion, and brand metrics without reducing creativity to short-term click optimisation)
- Asset governance (rights management, disclosure standards, provenance tracking, and auditability)

This reframes the organisation as a creative system rather than a collection of creative individuals. It also clarifies why AI may advantage smaller and agile firms: where evaluation routines and governance are well designed, AI lowers production costs and increases responsiveness. However, absent strong organisational capabilities, the same tools can create operational noise—many outputs, weak coherence, and diluted brand meaning.

6.5 Governance and legitimacy as constitutive, not peripheral

A major contribution of the synthesis is the demonstration that governance is not an external constraint applied after creativity occurs. Rather, governance is constitutive of creative value in AI-mediated marketing. In practice, questions of privacy, transparency, bias, authenticity, and authorship do not merely affect legal exposure; they shape consumer perceptions of trust and the symbolic legitimacy of brand communications. In AI-enabled contexts, the conditions under which content is created and targeted become part of the meaning of the message itself.

This is especially important because AI systems rely on data and probabilistic inference, which raises legitimacy concerns across the creative pipeline: the origins of training data, the fairness of targeting logic, the transparency of synthetic content, and the ethics of personalisation. The findings therefore advance the argument that market performance depends not only on the novelty and usefulness of creative outputs, but also on perceived legitimacy—whether audiences regard the communication as credible, respectful, and authentic. This responds directly to futures-oriented concerns about disruption, trust, and agency (McDowall, 2018), and reinforces the need to theorise governance as a moderator of creative effectiveness rather than an afterthought.

6.6 Theoretical contribution: creativity as a socio-technical, strategic system

Collectively, these insights extend marketing creativity theory in three ways.

First, the synthesis advances a system-level ontology of creativity, treating creative outcomes as emergent from interactions between AI capability, human judgement, organisational design, and governance. This moves beyond individual-centric creativity models and offers a more realistic account of how creativity is produced within contemporary marketing organisations (Ameen et al., 2022).

Second, it clarifies a strategic mechanism through which AI influences competitive advantage: AI expands ideation, but differentiation increasingly depends on evaluative and narrative capabilities that are unevenly distributed across firms (Pagani & Wind, 2024). As such, AI may raise the average quality of market communications while simultaneously intensifying competition around brand distinctiveness, cultural resonance, and trust.

Third, it positions legitimacy and ethics as central moderators of creative value creation, demonstrating that governance mechanisms shape not only compliance but also market acceptance. This directly addresses concerns raised in innovation and futures literature regarding the social viability of AI-driven marketing systems and the continuing centrality of human agency in navigating ambiguity and normative trade-offs (McDowall, 2018).

In summary, the integrated discussion indicates that AI is best theorised not as a creative agent that replaces human creativity, but as a general-purpose creative infrastructure that changes the economics, speed, and scope of creative work. Under this view, marketing creativity increasingly resides in human judgement, organisational design, and ethical stewardship, the capabilities that convert generative abundance into meaningful, differentiated, and legitimate brand communication.

7 Conclusion

This study set out to address key gaps in the literature by examining how artificial intelligence is reshaping creativity in marketing and advertising, with particular attention to human–AI collaboration and the strategic, organisational, and ethical implications of AI-driven disruption. Drawing on a theory-driven thematic synthesis of thirteen peer-reviewed journal articles and academic theses, the study provides an integrated explanation of AI-enabled creativity as a socio-technical and strategic system, rather than as a purely technological or efficiency-oriented phenomenon.

In response to the first research question, the findings demonstrate that AI is fundamentally reconfiguring, rather than replacing, marketing creativity. Generative and analytical AI tools significantly expand ideation capacity by enabling rapid production and exploration of creative alternatives. However, this expansion shifts the locus of creative value away from idea generation towards judgement, evaluation, and sensemaking. Creativity in AI-augmented contexts is therefore less concerned with producing ideas and more concerned with selecting, shaping, and legitimising them within brand, cultural, and market constraints. This challenges traditional views that locate creativity primarily in individual inspiration or artistic talent, and instead positions creativity as an outcome of structured interaction between humans, intelligent systems, and organisational arrangements.

Addressing the second research question, the analysis highlights the central role of human–AI collaboration design in shaping creative outcomes. Across the literature, effective collaboration is characterised by clear role allocation, with AI supporting exploration, variation, and experimentation, while humans retain authority over framing, evaluation, and narrative coherence. Human creative agency emerges not as a residual presence but as a necessary mediating mechanism that translates AI capability into meaningful marketing outcomes. Where such agency is weakened, AI-generated outputs tend to converge towards generic forms, undermining differentiation and strategic relevance.

In relation to the third research question, the study shows that AI-driven creative disruption has uneven strategic and organisational effects. Firms with stronger evaluative capabilities, higher levels of AI literacy, and clearer strategic intent are better positioned to convert AI-generated ideas into effective marketing outputs. While AI can act as a catalyst for creative disruption—particularly for smaller and more agile firms—such outcomes depend on complementary investments in skills, organisational learning, and business model innovation. Ethical and governance considerations also play a central role. Issues of privacy, transparency, authenticity, and legitimacy actively shape how AI-enabled creativity is perceived and accepted in the marketplace, forming part of the creative system rather than functioning as external constraints.

Overall, the study advances marketing creativity theory by shifting attention from individual-centric models towards system-oriented frameworks that integrate technological, organisational, and ethical dimensions. Practically, the findings suggest that organisations seeking to benefit from AI should prioritise evaluative and strategic capability development, design human–AI collaboration thoughtfully and embed ethical governance within creative workflows. In conclusion, artificial intelligence represents not the end of marketing creativity, but its reconfiguration. Creativity remains central to marketing success, increasingly residing in human judgement, organisational design, and ethical stewardship within AI-enabled marketing systems.

8 Ethics Declaration

This study employed a theory-driven thematic synthesis of publicly available, peer-reviewed academic literature and published doctoral theses. No primary data collection involving human participants, animals, or sensitive personal data was conducted. As the research draws exclusively on secondary sources already in the public domain, formal ethical clearance from an institutional review board was not required. All sources have been appropriately cited and referenced in accordance with standard academic practice.

9 AI Declaration

Artificial intelligence tools were used in the development of this paper for research assistance, writing support, and image generation. Specifically, an AI-powered conversational assistant (Kimi, developed by Moonshot AI) was used to support literature synthesis, refine argumentation, and assist with the drafting and editing of manuscript sections. The AI tool was employed as a collaborative writing aid, with the author retaining full responsibility for all analytical judgements, theoretical interpretations, and final editorial decisions. All substantive claims, propositions, and conclusions were critically evaluated and validated by the author prior to inclusion. Additionally, AI-powered image generation tool was used to produce visual material included in this paper. This image was generated to illustrate conceptual frameworks and was reviewed and approved by the authors for accuracy and relevance to the research content. The authors acknowledges that AI-generated outputs were treated as provisional drafts requiring human oversight, verification, and refinement. All cited sources were independently checked against original publications, and the author assumes full accountability for the intellectual integrity, accuracy, and originality of the final submitted work.

References

AbouElgheit, E. (2025). Generative AI as a disruptive innovation: Implications for marketing strategic transformations. *Foresight and STI Governance*, *19*(1), 6–15.

Ameen, N., Sharma, G. D., Tarba, S. Y., Rao, A., & Chopra, R. (2022). Toward advancing theory on creativity in marketing and artificial intelligence. *Psychology & Marketing*.

Albakry, N. S., Hashim, M. E. A., Harun, M. F., & Puandi, M. F. (2025). The role of artificial intelligence in creative design for advertising and digital content in educational contexts. *Semarak International Journal of Creative Art and Design*, *4*(1), 12–23.

Davenport, T. H., Guha, A., Grewal, D., & Bressgott, T. (2023). How artificial intelligence will change the future of marketing. *Journal of the Academy of Marketing Science*.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, *13*(3), 319–340.

Huang, M. H., & Rust, R. T. (2021). A strategic framework for artificial intelligence in marketing. *Journal of the Academy of Marketing Science*, *49*(1), 30–50.

Lee, H., & Cho, C. H. (2020). Digital advertising: Present and future prospects. *International Journal of Advertising*, *39*(3), 332–341.

McDowall, J. S. (2018). *The future of marketing: An investigation into disruption and innovation* (Doctoral thesis). University of Waikato.

Pagani, M., & Wind, J. (2024). Unlocking marketing creativity using artificial intelligence. *Journal of Interactive Marketing*, *60*(1).

Păvăloaia, V.-D., & Necula, S.-C. (2023). Artificial intelligence as a disruptive technology—A systematic literature review. *Electronics*, *12*(5), 1102.

Pradeep, A. K., Appel, A., & Sthanunathan, S. (2018). *AI for marketing and product innovation*. Wiley.

Rogers, E. M. (1962). *Diffusion of innovations*. Free Press.

Rust, R. T. (2020). The future of marketing. *International Journal of Research in Marketing*, *37*(1), 15–26.

Stone, M., Aravopoulou, E., Ekinci, Y., Evans, G., Hobbs, M., Labib, A., et al. (2020). Artificial intelligence in strategic marketing decision-making: A research agenda. *The Bottom Line*, *33*(2), 183–200.

Zaki, M. (2019). Digital transformation: Harnessing digital technologies for the next generation of services. *Journal of Services Marketing*, *33*(4), 429–435.

Making Knowledge Visible: The Role of Process Mapping in Enabling Knowledge Management in Industrial SMEs

Mariavittoria Marinucci[1], Giovanni Schiuma[2], Roberto Sala[1], Fabiana Pirola[1], Fabio Floreani[3]

[1] University of Bergamo, Bergamo, Italy
[2] LUM University, Casamassima, Italy
[3] Consorzio Intellimech, Italy

Abstract

This paper investigates how knowledge management (KM) practices manifest in industrial small and medium-sized enterprises (SMEs) and explores the role of process mapping as a mechanism for enabling the codification, sharing, and reuse of knowledge. While prior literature has extensively examined KM and process management as distinct streams, limited attention has been paid to their integration, especially within SMEs characterized by informal structures and fragmented knowledge practices. Adopting a qualitative, exploratory research design, the study combines a structured literature review with an empirical investigation based on semi-structured interviews conducted with a cluster of seven industrial SMEs. The literature review highlights three consistent findings: KM and Business Process Management (BPM) are mutually reinforcing; digitalization enhances the visibility and traceability of processes. However, these streams remain largely disconnected, revealing a gap in understanding how knowledge becomes operationally embedded within processes. The empirical findings show that SMEs demonstrate strong capabilities in knowledge creation and in increasing the adoption of digital tools, but struggle to translate this potential into structured, reusable organizational knowledge. Key KM dimensions such as reuse, decision traceability, and flow clarity remain underdeveloped. Process mapping emerges as a critical yet underutilized enabler: although firms recognize its importance, it is often implemented in a fragmented or purely descriptive manner, without explicit linkage to KM objectives. The study offers three main contributions. First, it proposes an integrated perspective that connects KM and process-based approaches. It also provides empirical evidence of the misalignment between knowledge generation and its codification within SMEs and, finally, it highlights process mapping as a foundational infrastructure for transforming tacit and dispersed knowledge into structured organizational assets. The findings suggest that investments in digital technologies and artificial intelligence are insufficient to improve KM outcomes unless supported by clearly defined and shared process structures. Future research should further explore how process mapping, process mining, and AI can be jointly leveraged to enhance knowledge-intensive operations in SMEs.

Keywords: Knowledge Management, Process Mapping; Industrial SMEs, Artificial Intelligence.

Paper type: Academic Research Paper

1 Introduction

In the contemporary economy, knowledge is widely recognized as a strategic resource for competitiveness, innovation, and organizational adaptability. The foundational literature on knowledge management (KM) has clarified that knowledge does not coincide with mere information: Davenport define it as a “fluid mix of framed experience, values, contextual information, and expert insight”, while Ikujiro Nonaka shows that organizational knowledge is generated through a “continuous dialogue between tacit and explicit knowledge”(Nonaka, 1994). A particularly relevant aspect of this paper is that such knowledge does not remain confined to individuals or documents but is also embedded in “routines, processes, practices, and norms” (Davenport et al., n.d.). In other words, KM becomes truly observable when it is translated into organizational arrangements, operational practices, and process structures.

This perspective becomes even more significant in industrial contexts characterized by service-oriented logics. The transition from product-centered models to solution- and service-oriented models makes knowledge a primary driver of value creation: what matters is not only designing and producing effectively, but also capturing, interpreting, and reusing knowledge generated throughout the offering's lifecycle, from operations to after-sales activities. In this regard, studies have shown that knowledge management is connected to organizational performance, highlighting that KM should be understood as a mechanism for enhancing the value organizations create (Schiuma et al., 2012a). In parallel, in service firms, the principle of “efficiency-of-use” tends to be replaced by “availability-to-use”: knowledge generates value to the extent that it is available, accessible, and reusable within processes. In recent years, this discussion has been further enriched by digital transformation, Industry 4.0, and the adoption of artificial intelligence (AI) tools. The literature shows that KM,

digitalization, and Industry 4.0 are converging into an increasingly structured research field (De Bem Machado et al., 2022). From an AI perspective, Jarrahi et al., (2023) underline that intelligent systems can support the main dimensions of KM—knowledge creation, storage, retrieval, sharing, and application—but only when they are effectively integrated with people, infrastructures, and organizational processes. This leads to a first key point: the presence of digital or AI tools does not automatically translate into effective knowledge management.

This issue is particularly evident in small and medium-sized industrial enterprises (SMEs). A well-established body of literature shows that KM practices in SMEs are often informal, fragmented, highly dependent on individual experience, and rarely supported by formalized systems and processes (Cerchione et al., 2016, 2015). More recent studies confirm that even amid increasing digitalization, SMEs tend to adopt traditional, isolated tools without developing a coherent knowledge management architecture (Castagna et al., 2020). Furthermore, the review by Durst et al., (2023) highlights that the field of KM in SMEs is rapidly expanding yet remains fragmented across loosely connected research streams, including knowledge sharing, capabilities, innovation, entrepreneurship, and digitalization. In this context, knowledge continues to be generated "in practice," but struggles to be codified, traced, and systematically reused. It is precisely at this point that process mapping becomes relevant. If, as (Davenport et al., n.d.) argue, knowledge is embedded in routines and processes, then process mapping is not merely a tool for operational efficiency, but also a mechanism for making knowledge visible within the organization. Recent literature on data-driven techniques shows that mapping actual workflows enables the identification of bottlenecks, deviations, tacit knowledge embedded in practices, and discrepancies between designed and executed processes (Lorenz et al., 2021). Even more critically, Wijnhoven et al., (2023) demonstrate that, in SMEs, the understanding and reconstruction of processes are particularly challenging due to the absence of clear de jure models and the ambiguity of information system semantics. This makes process mapping not a methodological accessory, but a potential enabling condition for knowledge codification, decision traceability, sharing, and reuse in service-intensive contexts.

Despite these advancements, a significant gap remains in the literature. Contributions on KM in SMEs have mainly focused on practices, barriers, and systems. However, there remains a lack of explicit integration among these streams that explains how knowledge is embedded and made manageable within operational and service processes in industrial SMEs, and what role process mapping plays in the transition from experiential, dispersed knowledge to codified, reusable organizational knowledge. This paper positions itself precisely at this theoretical and empirical intersection. The research question guiding the study is as follows: how do knowledge management practices manifest in industrial SMEs, and what role does process mapping play in enabling the codification, sharing, and reuse of knowledge within operational and service processes? The expected contribution is threefold. From a theoretical perspective, the paper connects KM and process perspectives into a unified interpretative framework. From an empirical standpoint, it provides qualitative evidence from a cluster of industrial SMEs, a context that remains underexplored from an integrated perspective. From a methodological perspective, it combines a structured literature review, structured interviews, and coded data analysis to examine the extent to which knowledge remains informal or, conversely, is translated into observable process structures supported by digital tools. Consistent with this positioning, the methodology of the paper is articulated into two complementary phases. The first consists of a structured literature review based on Scopus queries combining the terms knowledge management, SMEs, service, process mapping/process mining, digital transformation, and AI, followed by screening and thematic analysis of the selected contributions. The second phase is a qualitative empirical study based on structured interviews conducted with seven firms belonging to an industrial cluster; the responses were subsequently codified into a dataset and analyzed using Python through descriptive matrices, heatmaps, gap analysis, and lexical analysis. In this way, the paper integrates bibliographic and empirical evidence to investigate whether and how processes make knowledge truly visible, shareable, and reusable.

2 Methodology

This study adopts a qualitative and exploratory research design aimed at investigating the role of process mapping in enabling knowledge management practices within industrial SMEs. The methodological approach combines a structured analysis of the academic literature with an empirical investigation based on semi-structured interviews conducted within an industrial cluster. This dual approach allows for the integration of theoretical insights and empirical evidence, providing a comprehensive understanding of the phenomenon under study.

The first phase of the research consists of a structured literature review conducted using the Scopus database, which was selected due to its broad coverage of peer-reviewed journals and its relevance for management and engineering studies. The objective of this phase was to identify and systematize the main

contributions at the intersection of knowledge management, SMEs, servitization, process mapping, and digital transformation. The search query was designed to capture the multidimensional nature of the research topic and included the following keywords:

("knowledge management" OR "organizational knowledge") AND ("SMEs" OR "small AND medium enterprises") AND ("process mapping" OR "business process management" OR "process mining") AND ("servitization" OR "product-service systems" OR "service") AND ("digital transformation" OR "artificial intelligence") AND (LIMIT-TO (SUBJAREA , "BUSI") OR LIMIT-TO (SUBJAREA , "ENGI")) AND (LIMIT-TO (DOCTYPE , "ar")) AND (LIMIT-TO (LANGUAGE , "English"))

The search was conducted without restricting the initial time window, thereby including both foundational and recent contributions. The results were then filtered based on document type (articles and reviews), language (English), and relevance to the research objectives. Abstracts and full texts were screened to exclude contributions that were not directly related to knowledge management practices or that did not provide insights into process-related dimensions.

The selected articles were analyzed using a thematic approach, focusing on four main streams: (i) knowledge management foundations and processes, (ii) KM in SMEs, (iii) servitization and service-oriented transformation, and (iv) process mapping and process mining. This analysis revealed a fragmentation of the literature across these streams and highlighted a lack of integration between knowledge management and process-based perspectives, particularly in SME contexts. The literature review thus served both to ground the research theoretically and to identify the gap addressed by the empirical investigation.

The second phase of the study is based on an empirical investigation conducted within a cluster of industrial SMEs. The cluster represents a collaborative environment involving firms and research institutions, aimed at fostering knowledge exchange and supporting innovation processes. The companies involved operate in manufacturing sectors characterized by high technical complexity and increasing service orientation, making them particularly relevant for the study of knowledge management in product-service contexts.

The selection of firms was driven by their participation in the cluster initiative and by their willingness to engage in a structured reflection on knowledge management practices. While the firms differ in size, organizational structure, and technological maturity, they share characteristics typical of SMEs, including limited process formalization and a strong reliance on individual expertise. Data were collected through semi-structured interviews conducted with key organizational actors, including managers and technical experts involved in operational and decision-making processes. The interview protocol was designed to explore how knowledge is created, codified, shared, and used within the organization, as well as the role of processes and digital tools in supporting these activities. Attention was devoted to identifying critical issues and perceived barriers in managing knowledge, as well as to understanding the role of process mapping practices, where present.

Following data collection, the interview responses were transcribed and systematically codified into a structured dataset. Each response was analyzed and mapped onto a set of analytical dimensions derived from the literature review, including knowledge creation, codification, sharing, use, reuse, process mapping, decision traceability, digital tools, and technology effectiveness. To enable comparison across firms, each dimension was evaluated on a three-point scale, with values indicating the degree of presence or maturity of the specific practice. This coding process transformed qualitative data into a semi-quantitative dataset, enabling the identification of patterns and differences across cases.

The resulting dataset was organized in a tabular format and subsequently analyzed using Python. The analysis included constructing a matrix representation of the data, visualizing it with heatmaps, calculating descriptive statistics, and developing gap analysis indicators to identify misalignments across key dimensions. In addition, a textual analysis of interview responses was conducted by extracting and processing the answers, enabling the identification of recurring themes and dominant concepts through word frequency analysis.

3 Literature Review

Recent literature converges on a key point: in industrial SMEs, process mapping is not merely a descriptive exercise, but a cognitive and organizational infrastructure that makes knowledge more visible, transferable, and reusable. When processes are explicitly represented through maps, "as-is/as-designed" models, roles, inputs/outputs, handoffs, exceptions, and document flows, tacit knowledge dispersed within operational practices is progressively translated into more stable organizational knowledge. In industrial SMEs, this role becomes even more critical, as value creation depends on integrating production processes, service processes, usage data, customer interactions, and decision-making activities across the entire lifecycle of the offering.

Digitalization and, more recently, artificial intelligence further amplify this role, but only when a sufficiently standardized and governed process foundation is in place.

3.1 KM, process orientation and SMEs

Knowledge does not generate value in an abstract way, but rather when it is embedded within the processes that support performance and value creation. Schiuma et al., (2012) emphasizes the link between knowledge management and performance improvement, while Schiuma et al., (2012b) interpret knowledge processes as mechanisms of value creation. This perspective is particularly consistent with industrial SMEs, where knowledge is often embedded in people, routines, and operational relationships rather than in formalized systems. Indeed, reviews on KM in SMEs clearly show that such firms are characterized by resource constraints, low levels of formalization, and a strong reliance on human and contextual knowledge—factors that make a lightweight yet process-embedded infrastructure crucial for stabilizing knowledge (Centobelli et al., 2019; Cerchione et al., 2016; Durst et al., 2023). More recent studies on SMEs also demonstrate that KM practices influence both internal performance and external relational capabilities, but their effectiveness depends on their integration into organizational processes rather than on their isolated existence(Rošulj et al., 2024).

From this, an important theoretical implication emerges: process mapping is valuable for KM because it allows knowledge to be "attached" to a sequence of activities, actors, decision rules, and documents. In other words, it makes organizational knowledge identifiable within the workflow and, therefore, manageable. Bitkowska, (2020), in an empirical study of Polish firms, explicitly demonstrates that Business Process Management positively influences knowledge management processes. The statement is clear: "Business Process Management has a positive influence on Knowledge Management processes"(Bitkowska, 2020). This finding is consistent with (Atkočiūnienė et al., 2022), who demonstrate that KM processes positively affect business processes and that business processes mediate the relationship between KM and value creation. The relationship, therefore, is not linear but mediated: knowledge generates value when it flows through processes that are observable, improvable, and measurable.

3.2 Process mapping as an enabler of coding, traceability, sharing and reuse

The first mechanism is codification. Mapping a process means defining activities, inputs, outputs, responsibilities, control points, exceptions, and required knowledge; in other words, it involves transforming what was previously situated, implicit, and fragmented into a shared organizational artifact. (Bitkowska, 2018) even interprets the Business Process Management Centre of Excellence as a source of knowledge, that is, as a structure that collects, organizes, and disseminates process-related knowledge. Even more explicitly, Sarnikar and Deokar (2017), from a design science perspective, show that process-based knowledge management systems support knowledge-intensive processes precisely because they link knowledge needs to the context in which the process is performed. For SMEs, this translates into reduced dependence on individual experts, increased replicability of effective practices, and faster onboarding, training, and standardization. The second mechanism is traceability. Here, the literature on event logs and process mining complements traditional process mapping approaches. Lorenz et al., (2021), in a make-to-stock manufacturing context, observe that traditional mapping methods "only provide static snapshots of process flows" (p. 4870), whereas process mining enables the dynamic mapping of real processes and the identification of deviations, bottlenecks, and non-value-adding activities. Wijnhoven et al., (2023), in an SME case study, further show that incomplete de jure models hinder the identification of workarounds and that human interpretation of results is necessary to activate triple-loop organizational learning. This theoretical point is highly relevant to this study: traceability does not depend solely on data but on the alignment between expected and actual processes. In the absence of a robust reference map, even digital analysis loses explanatory power. The third mechanism is sharing. Process maps function as boundary objects, providing a common language across functions and roles and, in firms, also across production, after-sales, maintenance, and customer interface activities. In this sense, mapping does not simply clarify "who does what," but makes knowledge visible, discussable, and negotiable by highlighting where it is generated, validated, interrupted, or lost. The review by Thiede et al., (2018) is particularly relevant in this regard, as it shows that organizational process mining has mainly been applied in single-system and single-organization contexts, while cross-system, cross-organizational, and service-oriented analyses remain underexplored. This gap is precisely what makes the case of industrial SMEs integrating internal processes and customer services particularly relevant. The fourth mechanism is reuse. When knowledge is codified and traced within processes, it can be retrieved and applied in subsequent contexts, such as anomaly management, training, process redesign, standardization, continuous improvement, and customer-oriented decision-making. Atkočiūnienė et al., (2022) show that KM processes should not be treated separately but integrated into business processes to

maximize value creation. In other words, reuse is not simply the retrieval of information, but the organization's ability to reintegrate validated knowledge into subsequent operational and decision-making cycles

4 Theoretical Gap

The analyzed literature highlights three robust findings. Business Process Management (BPM) and Knowledge Management (KM) are mutually reinforcing: processes structure knowledge, and knowledge, in turn, improves processes (Atkočiūnienė et al., 2022; Bitkowska, 2020; Schiuma et al., 2012b). Moreover, digitalization expands the potential of this relationship by transforming process mapping into a dynamic analysis of real processes (Goel et al., 2024; Lorenz et al., 2021; Wijnhoven et al., 2023). Finally, in servitization contexts, the relevance of process mapping increases because critical knowledge is distributed across production and service activities, customer interactions, and usage data (Castagna et al., 2020; Kohtamäki et al., 2020; Pirola et al., 2020).

However, these streams remain largely disconnected. The literature on BPM and KM rarely addresses the specific context of service-oriented industrial SMEs.

This leads to a clear theoretical gap: there is a lack of an integrated framework that explains how and under which conditions process mapping concretely enables KM outcomes—such as codification, traceability, sharing, and reuse.

5 Analysis

The empirical analysis is based on a cluster of seven industrial SMEs operating in manufacturing sectors characterized by high levels of technical complexity and an increasing orientation towards service-based offerings. The firms are part of a structured research initiative aimed at exploring knowledge management practices and digital transformation trajectories within industrial ecosystems. To ensure confidentiality, all companies have been anonymized. Despite differences in organizational structure and market positioning, they share typical SME characteristics, including limited process formalization, strong reliance on individual expertise, and the adoption of digital tools that are not always fully integrated into organizational practices.

5.1 Data structure and analytical framework

Data were collected through semi-structured interviews and subsequently codified into a structured dataset. Each interview response was mapped onto a set of analytical dimensions derived from the knowledge management literature, including knowledge creation, codification, sharing, use, reuse, process mapping, decision traceability, digital tools, and technology effectiveness. Each dimension was evaluated using a three-point scale, enabling both cross-case comparison and aggregation of results. The dataset was then analyzed using Python-based techniques, including matrix visualization, descriptive statistics, gap analysis, and textual analysis of interview responses, allowing for a combined quantitative and qualitative interpretation of the findings.

5.2 Cross-firm analysis of KM and process dimensions

The matrix visualization, in Figure 1, provides a comprehensive overview of how knowledge management and process-related dimensions are distributed across the analyzed firms. The results reveal a clear heterogeneity in the adoption of process-oriented practices. While some firms exhibit relatively structured approaches to process mapping, others show limited formalization, confirming the uneven diffusion of process-based management within SMEs.

At the same time, dimensions related to knowledge creation and digital tools tend to display consistently higher values across firms. This suggests that organizations can generate knowledge through their operational activities and have access to technological infrastructures that could potentially support knowledge management processes. Hoever, this potential is not systematically translated into structured organizational knowledge. In contrast, dimensions such as reuse, decision traceability, and flow clarity show lower and more fragmented scores. This indicates that knowledge, although generated, is rarely formalized, tracked, or reused across organizational processes. As a result, knowledge remains largely embedded in individual experience and localized practices rather than being incorporated into shared organizational structures. These findings support the idea that, in the absence of explicit process integration, knowledge management remains informal and difficult to scale.

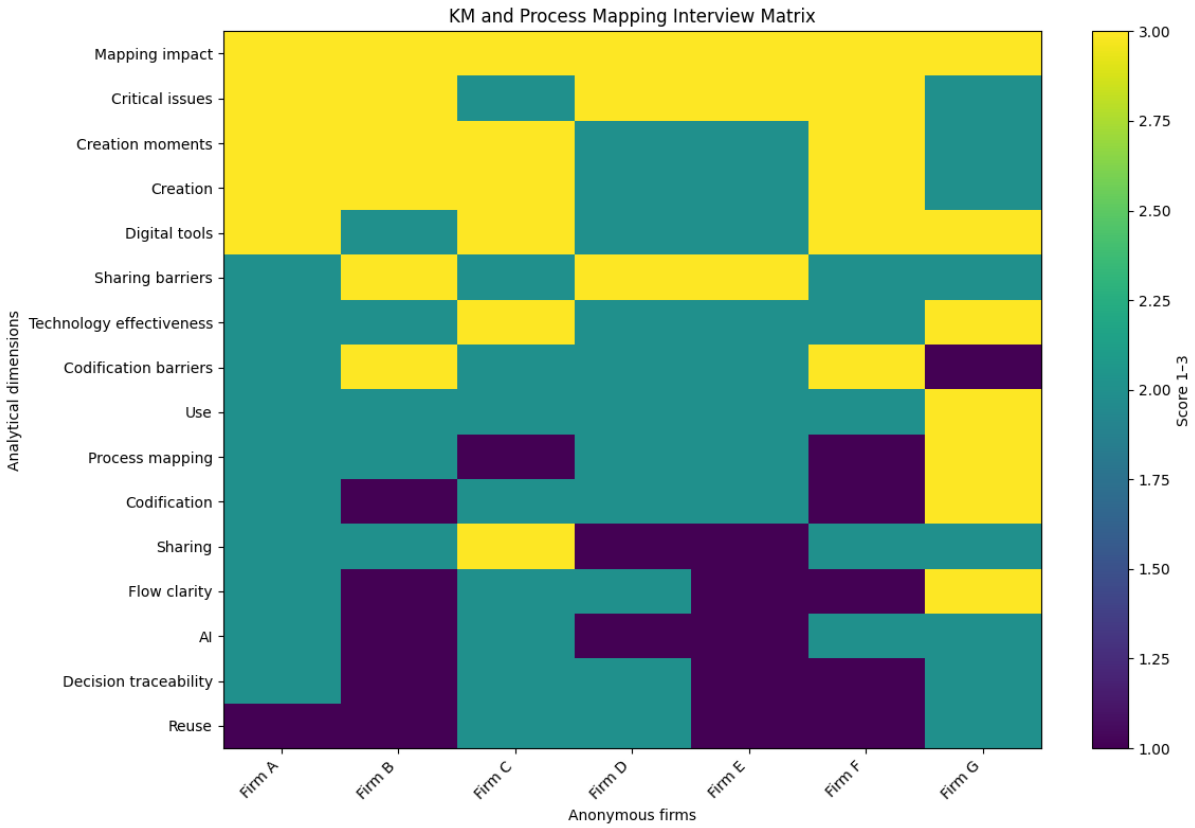


Figure 1 - Heatmap of KM dimensions across firms

5.3 Descriptive analysis of KM dimensions

The analysis of average scores further clarifies the overall maturity of knowledge management practices across the investigated dimensions. Higher values are observed in dimensions such as mapping impact, knowledge creation, and digital tools, suggesting that firms recognize the importance of knowledge and its role in addressing operational challenges. In particular, the high score associated with mapping impact indicates that organizations perceive process mapping as a potentially valuable mechanism for improving knowledge management, even when its practical implementation remains limited.

However, lower scores emerge for dimensions such as reuse, decision traceability, and flow clarity. This highlights a structural limitation in the way knowledge is managed within these organizations. While knowledge is generated and partially supported by digital technologies, it is not effectively embedded into processes that enable systematic reuse and informed decision-making. The relatively moderate score of process mapping itself further reinforces this interpretation, suggesting that although its importance is acknowledged, it is not consistently implemented or integrated with other knowledge management practices.

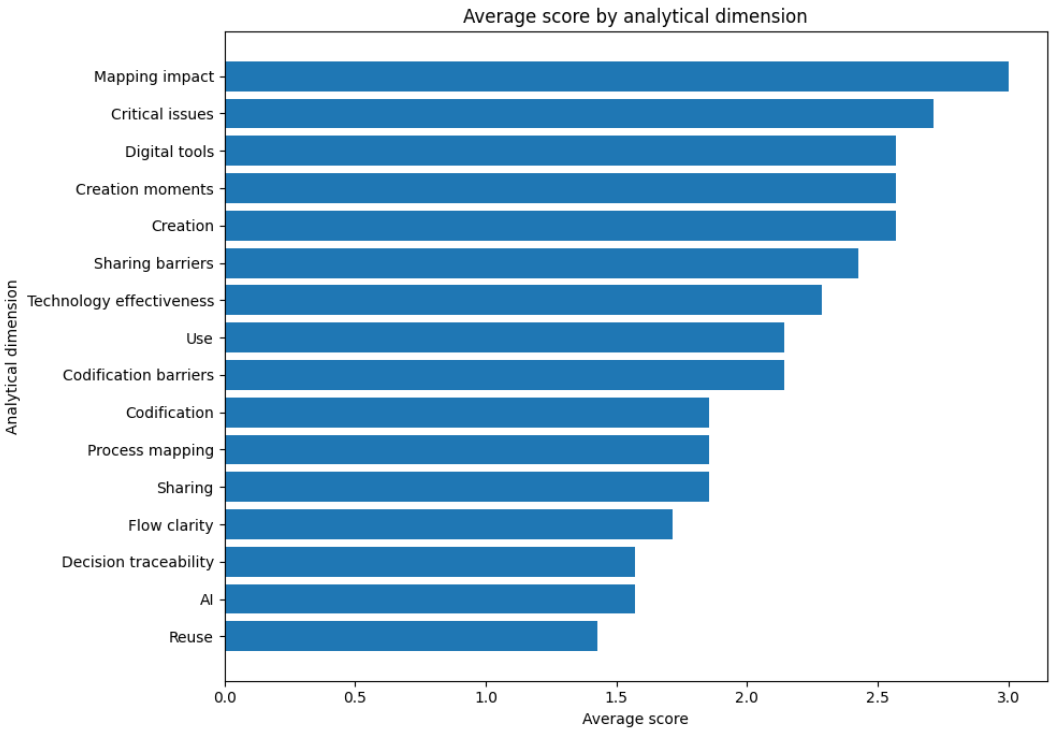


Figure 2 - Average KM maturity across analytical dimensions

5.4 Gap analysis across firms

The gap analysis provides deeper insight into the structural misalignments that characterize knowledge management practices within the analyzed firms. A consistent pattern emerges when comparing knowledge creation with codification. In most cases, the ability to generate knowledge exceeds the capacity to formalize it, indicating that knowledge remains tied to operational experience rather than being translated into structured and reusable organizational assets. A similar misalignment is observed between the presence of digital tools and their effective use. While firms report having technological solutions available, these are not always fully exploited to support knowledge management processes. This confirms that the mere presence of digital infrastructure does not guarantee effective knowledge integration, particularly when it is not embedded within clearly defined processes. Perhaps the most significant gap concerns the relationship between process mapping and knowledge reuse. Even when process mapping practices are present, they do not systematically lead to improved knowledge reuse. This suggests that process mapping, when not explicitly connected to knowledge management objectives, remains a descriptive or operational tool rather than a mechanism for enabling organizational learning.

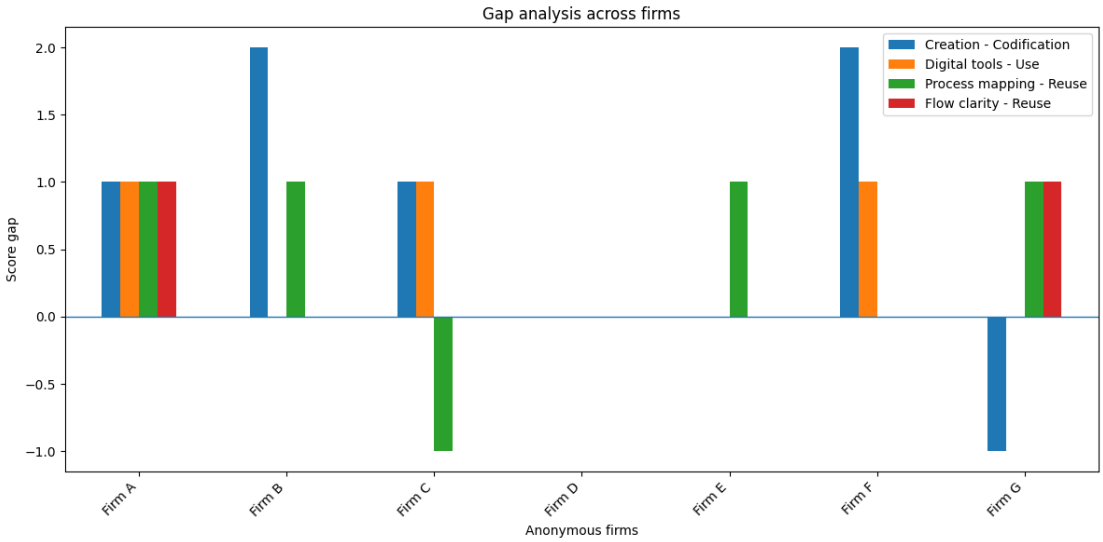


Figure 3 - Cross-firm gap analysis of KM

6 Conclusion

This study set out to investigate how knowledge management practices are enacted within industrial SMEs and to explore the role of process mapping in enabling the codification, sharing, and reuse of organizational knowledge, particularly in service-oriented contexts. By combining a structured literature review with an empirical analysis based on interviews, the paper provides both theoretical and practical insights into the relationship between knowledge and processes. The findings highlight a consistent pattern across the analyzed firms. Knowledge is actively generated through operational activities and supported, at least in part, by digital tools. However, the absence of structured, integrated processes hinders the transformation of this knowledge into reusable organizational assets. Dimensions such as reuse, decision traceability, and flow clarity remain underdeveloped, confirming that knowledge often remains embedded in individual experience rather than being systematically managed.

Process mapping emerges as a critical, yet underutilized, mechanism for addressing this gap. While firms recognize its importance, implementation is often partial and not fully integrated with knowledge management practices. As a result, the potential of process-based approaches to support learning, decision-making, and service delivery remains largely untapped.

From a theoretical perspective, this study contributes to the literature by integrating knowledge management and process-based views, highlighting the central role of processes in making knowledge visible, transferable, and reusable. From a managerial standpoint, the results suggest that investments in digital tools or AI are insufficient without a structured approach to process mapping and knowledge integration.

The study is subject to limitations, primarily related to the qualitative nature of the analysis and the limited number of firms involved. Future research could extend the sample, adopt longitudinal approaches, or further explore the integration between process mapping, digital technologies, and AI in knowledge-intensive and service-oriented contexts.

7 Ethics Declaration

The authors declare that this study did not involve personal sensitive data, or experiments requiring formal ethical approval. The empirical investigation was conducted through voluntary semi-structured interviews with company representatives. All the collected information was anonymized and treated confidentially. The study complies with the ethical standards of academic research.

8 AI Declaration

The authors declare that artificial intelligence (AI) tools were used exclusively to support language revision and text editing during the preparation of this manuscript. All intellectual contributions, interpretations, methodological choices, and final revisions remain the sole responsibility of the authors.

References

Atkočiūnienė, Z.O., Gribovskis, J., Raudeliūnienė, J., 2022. Influence of Knowledge Management on Business Processes: Value-Added and Sustainability Perspectives. Sustainability 15, 68. https://doi.org/10.3390/su15010068

Bitkowska, A., 2020. The relationship between Business Process Management and Knowledge Management - selected aspects from a study of companies in Poland. JEMI 16, 169–193. https://doi.org/10.7341/20201616

Bitkowska, A., 2018. BUSINESS PROCESS MANAGEMENT CENTRE OF EXCELLENCE AS A SOURCE OF KNOWLEDGE. Business, Management and Education 16, 121–132. https://doi.org/10.3846/bme.2018.2190

Castagna, F., Centobelli, P., Cerchione, R., Esposito, E., Oropallo, E., Passaro, R., 2020. Customer Knowledge Management in SMEs Facing Digital Transformation. Sustainability 12, 3899. https://doi.org/10.3390/su12093899

Centobelli, P., Cerchione, R., Esposito, E., 2019. Efficiency and effectiveness of knowledge management systems in SMEs. Production Planning & Control 30, 779–791. https://doi.org/10.1080/09537287.2019.1582818

Cerchione, R., Esposito, E., Spadaro, M., 2015. The Spread of Knowledge Management in SMEs: A Scenario in Evolution. Sustainability 7, 10210–10232. https://doi.org/10.3390/su70810210

Cerchione, R., Esposito, E., Spadaro, M.R., 2016. A literature review on knowledge management in SMEs. Knowledge Management Research & Practice 14, 169–177. https://doi.org/10.1057/kmrp.2015.12

Davenport, T.H., Prusak, L., Webber, A., n.d. Working Knowledge: How Organizations Manage What They Know.

De Bem Machado, A., Secinaro, S., Calandra, D., Lanzalonga, F., 2022. Knowledge management and digital transformation for Industry 4.0: a structured literature review. Knowledge Management Research & Practice 20, 320–338. https://doi.org/10.1080/14778238.2021.2015261

Durst, S., Edvardsson, I.R., Foli, S., 2023. Knowledge management in SMEs: a follow-up literature review. JKM 27, 25–58. https://doi.org/10.1108/JKM-04-2022-0325

Goel, K., Martin, N., Ter Hofstede, A., 2024. Demystifying data governance for process mining: Insights from a Delphi study. Information & Management 61, 103973. https://doi.org/10.1016/j.im.2024.103973

Jarrahi, M.H., Askay, D., Eshraghi, A., Smith, P., 2023. Artificial intelligence and knowledge management: A partnership between human and AI. Business Horizons 66, 87–99. https://doi.org/10.1016/j.bushor.2022.03.002

Kohtamäki, M., Parida, V., Patel, P.C., Gebauer, H., 2020. The relationship between digitalization and servitization: The role of servitization in capturing the financial potential of digitalization. Technological Forecasting and Social Change 151, 119804. https://doi.org/10.1016/j.techfore.2019.119804

Lorenz, R., Senoner, J., Sihn, W., Netland, T., 2021. Using process mining to improve productivity in make-to-stock manufacturing. International Journal of Production Research 59, 4869–4880. https://doi.org/10.1080/00207543.2021.1906460

Nonaka, I., 1994. A Dynamic Theory of Organizational Knowledge Creation.

Pirola, F., Boucher, X., Wiesner, S., Pezzotta, G., 2020. Digital technologies in product-service systems: a literature review and a research agenda. Computers in Industry 123, 103301. https://doi.org/10.1016/j.compind.2020.103301

Rošulj, D., Petrović, D.Č., Arsić, S.M., 2024. Knowledge Management in Serbian SMEs: Key Factors of Influence on Internal and External Business Performances. Sustainability 16, 797. https://doi.org/10.3390/su16020797

Sarnikar, S., Deokar, A.V., 2017. A design approach for process-based knowledge management systems. JKM 21, 693–717. https://doi.org/10.1108/JKM-09-2016-0376

Schiuma, G., Carlucci, D., Lerro, A., 2012a. Managing knowledge processes for value creation. VINE 42, 4–14. https://doi.org/10.1108/03055721211207815

Schiuma, G., Carlucci, D., Lerro, A., 2012b. Managing knowledge processes for value creation. VINE 42, 4–14. https://doi.org/10.1108/03055721211207815

Thiede, M., Fuerstenau, D., Bezerra Barquet, A.P., 2018. How is process mining technology used by organizations? A systematic literature review of empirical studies. BPMJ 24, 900–922. https://doi.org/10.1108/BPMJ-06-2017-0148

Wijnhoven, F., Hoffmann, P., Bemthuis, R., Boksebeld, J., 2023. Using process mining for workarounds analysis in context: Learning from a small and medium-sized company case. International Journal of Information Management Data Insights 3, 100163. https://doi.org/10.1016/j.jjimei.2023.100163

AI-Supported Decision-Making in Medical Device Procurement: From Ad Hoc Criteria to Meta-Criteria and Value Configurations

Mattia Barsanti, Salvatore Tallarico, Luisa Pellegrini, Riccardo Dulmin, Simone Lazzini, Marco Praticò
University of Pisa, Pisa, Italy

Abstract

Medical device (MD) procurement is a strategic area of non-clinical decision-making in healthcare, as purchasing choices affect not only cost containment but also quality of care, patient safety, and long-term health system sustainability. In competitive tenders, however, evaluation criteria are often highly heterogeneous and device specific. This limits the possibility of systematically comparing procurement decisions across tenders and of understanding the evaluative logics that guide final decisions. Within this context, artificial intelligence (AI) may support procurement not by replacing human judgement, but by structuring fragmented information, identifying recurrent value dimensions, and enabling a more reflexive comparison between AI-generated outputs and observed human decisions.

This study investigates how AI can support decision-making in MD procurement. It focuses on tender documents managed by a regional central purchasing body in Italy and develops a research design structured into four phases: tender document analysis and data extraction; recoding specific-criteria into higher-level meta-criteria; construction of AI-based value configurations; and comparison of these configurations with observed procurement decisions. The paper reports the preliminary results of the first two phases.

The paper reports the preliminary results of the first two phases. As regards the first phase, approximately 500 tenders were identified. After excluding non-MD tenders, price-only procedures, incomplete cases, and non-interpretable documents, the final dataset consisted of 172 tender documents. Preliminary analysis identified three recurring document formats and corresponding extraction schemas.

As regards the second phase, the first results show that device-specific evaluation criteria can be recoded into broader meta-criteria, such as versatility of use, patient safety, operator safety, interoperability, and degree of clinical certainty. These meta-criteria provide the analytical basis for comparing tenders involving different types of MDs and for integrating AI appropriately into procurement decision-making. By preserving traceability to the original criteria while enabling broader comparability, they support AI-based procurement governance and may contribute to accountability and organizational learning by supporting a more explicit justification of procurement decisions over time.

Keywords: Artificial Intelligence, MD, tender, public health, meta-criteria

Paper type: Academic Research Paper

1 Introduction and Theoretical Background

Medical device (MD) procurement represents a critical area of non-clinical decision-making in healthcare, with direct implications for cost containment, quality of care, and the long-term sustainability of health systems. Purchasing decisions – particularly those managed by central purchasing bodies through competitive tenders –

play a pivotal role in shaping not only economic efficiency but also the effectiveness and safety of care delivery. In this context, procurement processes are expected to incorporate a plurality of evaluation dimensions, including technical performance, clinical quality, lifecycle costs, maintenance requirements, training needs, and supplier-related characteristics (Hinrichs-Krapels et al., 2022; Miller et al., 2019).

Within this scenario, artificial intelligence (AI) emerges as a potentially transformative yet still underexplored tool. Research from supply chain management and public procurement suggests that AI (e.g., machine learning–based Natural Language Processing (NLP)) can support complex decision-making processes and, under controlled conditions, match or outperform traditional analytical approaches (Long et al., 2023). Specifically, this body of literature highlights the potential of AI to assist across several phases of the procurement process, including supplier evaluation, tender matching, data analytics, and decision support (Balkan & Akyuz, 2025; Siciliani et al., 2023; Allal-Chérif et al., 2020).

Empirical studies have begun to explore these applications in public procurement contexts. For instance, focusing on the data analytics phase, Nai et al. (2023) show that machine learning models can detect anomalous procurement patterns and identify tenders at higher risk of legal disputes. Similarly, in the decision-support phase, Acikalin et al. (2024) develop predictive models based on textual analysis of EU procurement notices,

demonstrating that the linguistic and structural quality of tender documentation can influence outcomes such as participation rates and cost overruns.

In addition, a key potential contribution of AI lies in its ability to reconstruct the implicit evaluative logic that has guided previous choices and hence generate outputs grounded in past decisions (e.g., rankings, scores, or recommendations) that can serve as benchmarks for the new decision Specifically, these AI-generated outputs can support decision-makers to identify to what extent current decisions are consistent with expectations and analyse where divergences between AI-based recommendations and actual human decisions occur. This may prompt reflection on why differences arise – whether due to contextual changes, evolving priorities, or hidden/unintended shifts in evaluative logic, subjective weightings, or inconsistencies in the interpretation of criteria – issues already recognized in procurement and supplier evaluation contexts (Rainy & Chowdhury, 2022; Kufile et al., 2022).

At the same time, AI can contribute to a more foundational challenge: the structuring of procurement information. Tender documents are often heterogeneous, unstructured, and poorly standardized, limiting their usability for systematic analysis. AI – particularly NLP based approaches – can extract, organize, and formalize information from these documents, transforming them into structured knowledge bases suitable for evaluation, comparison, and predictive modelling (Siciliani et al., 2023). This phase is not merely technical but strategic, as it enables more consistent application of evaluation criteria and supports the development of data-driven procurement practices.

Importantly, the adoption of AI also opens opportunities beyond decision support. By making evaluation criteria and decision logics more explicit, AI can help uncover implicit assessment frameworks, provide manufacturers with structured information derived from past tenders on the key evaluation dimensions, and thereby enable them to better align their offerings with evaluation logics, increasing their ability to compete effectively and succeed in future tenders.

Despite these opportunities, the application of AI to MD procurement remains limited and largely experimental (Guida et al., 2023), partly because the highly regulated and multidimensional nature of this context makes the concrete integration of AI more challenging than in other domains due to significant governance and regulatory challenges, including data quality, integration into existing workflows, algorithmic transparency, and accountability (Balkan & Akyuz, 2025; Khan et al., 2024). Robust oversight, continuous monitoring, and clear attribution of responsibility are essential to ensure trust and legitimacy in AI-supported decision-making, particularly in the public healthcare sector (Bosmans et al., 2021; Muehlematter et al., 2021; Angehrn et al., 2020).

In addition, limited attention has been paid to how AI systems can leverage historical tender data to generate predictive outputs meaningfully confronted with human procurement decisions. Rather than replacing human judgment, such comparisons allow identification of possible misalignments and support critical assessment of the robustness of procurement decisions (Siciliani et al., 2023).

To address these issues, this study proposes shifting the focus from *ad hoc* criteria tailored to specific MDs to higher-level evaluative dimensions, referred to as meta-criteria. For instance, meta-criteria abstract from device-specific technical attributes (e.g., balloon catheter size) and capture broader dimensions of value, such as clinical adequacy (e.g., Dimension, weight, and footprint). By abstracting from *ad hoc* criteria to meta-criteria, AI models can capture more stable and comparable evaluative logics across tenders and identify context-specific value configurations expected in terms of meta-criteria. By explicitly engaging with emerging literature on AI-based decision support in public procurement, this study conceptualizes predictive outputs not as substitutes for human judgment, but as reflexive tools that support critical assessment of procurement decisions and their robustness (Cernăzanu-Glăvan & Bulzan, 2025; Zorina & Bikir, 2025). By explicitly identifying where and why a decision diverges from expected value configurations, this approach strengthens accountability in public procurement (Aboelazm & Dganni, 2025; Guida et al., 2023).

The focus is on procurement processes managed by a central purchasing body at the regional level in Italy, a context characterized by variability in tender structures and evaluation practices (Atella & Decarolis, 2019; Borsoi et al., 2017). Within this setting, the study explores not only the comparison between AI-generated outputs and human decisions, but also the preliminary and critical phase of transforming unstructured tender documentation into a structured and analysable knowledge base. By doing so, it aims to assess how AI can enhance comparability, transparency, and consistency in procurement decisions, while also shedding light on the underlying evaluative logics that shape them.

At its current stage, however, the study presents only the initial outputs of this research agenda. These include the systematic analysis of tender documents, their classification, the extraction of evaluation criteria, an initial recoding into higher-level meta-criteria (at this stage only partially defined), and the preliminary

construction of a structured dataset. These phases represent the foundational phase necessary to enable subsequent AI-based modelling and decision-support analyses.

Accordingly, the broader research project is guided by the following research question:

RQ. How can AI support decision-making in MD procurement?

2 Methodology

The research project is structured around four distinct phases

1. Tender documents analysis and data extraction (see section 2.1),
2. Recoding of evaluation criteria into meta-criteria (see section 2.2),
3. Combination of meta-criteria to construct AI-based value configurations,
4. Comparison of these configurations with observed human procurement decisions.

Below, we detail the first two phases, which are currently underway.

2.1 Tender document analysis and data extraction

Tender documents are analysed to identify the evaluation criteria used in the procurement of MDs. The overall objective is to assess both the availability and the structure of the data contained within these documents, and to determine whether the information is suitable for training predictive models, including machine learning and neural network approaches, in line with the objectives of the research. This assessment considers not only the content of the documents – such as the presence of detailed and structured evaluation data – but also their volume, which is a critical factor for the development of reliable AI models. In this way, the analysis provides key insights into the quality, consistency, and adequacy of the dataset, while simultaneously informing the data extraction process.

Data extraction is performed using a generative AI chatbot based on a family of large language models. The accuracy and reliability of the extracted data are ensured through systematic manual validation, carried out by comparing the outputs with the original tender documents.

2.2 Recoding of evaluation criteria into meta-criteria

The meta-criteria were identified using a bottom-up approach. Initially, as the extracted data included not only evaluation criteria but also other features (such as suppliers name, manufacturer nationality, lot IDs, and assessment scores), only the evaluation criteria themselves were selected for further processing. These criteria were then grouped into meta-criteria designed to be applicable across different tenders, enabling greater generalisability. Meta-criteria are conceived as abstract categories that capture recurrent evaluative patterns across different procurement contexts.

To ensure their validity, the identified meta-criteria were compared with established criteria commonly found in recognised frameworks for MD evaluation in the literature (for example, EUnetHTA's Core Model), such as safety, usability, clinical appropriateness, and organisational impact. This comparison was intended to verify whether the meta-criteria were supported by the literature and thus robust for use in this context.

This approach allows the meta-criteria to be both grounded in the empirical data and aligned with established evaluative dimensions, ensuring both relevance to the specific procurement context and broader comparability across tenders.

3 Results

This section presents results by methodological phase.

3.1 Tender document analysis and data extraction

The analysis is based on tender documents retrieved from the official website of ESTAR (Regional Technical and Administrative Support Agency of Tuscany). The choice of this procurement body is particularly relevant due to the high level of centralisation of public procurement in Tuscany, which results in a single, unified online repository where all tender-related documentation is systematically published. This feature enables access to a large volume of documents and ensures data completeness across the entire procurement process, from tender specifications to final award decisions.

At the beginning of the research (January 2026), approximately 500 tenders classified under the furniture category were identified. Since this category includes heterogeneous items (e.g. pharmaceuticals, MD, office

furniture, and supplies), an initial filtering phase was required to isolate only MD-related tenders, reducing the dataset to approximately 250 cases.

The second selection stage consisted of excluding tenders based exclusively on price, as they do not provide multidimensional evaluation criteria. Additional exclusions concerned cases with incomplete evaluation data (e.g. absence of criterion-level scores) and documents whose structure did not allow reliable interpretation or comparison. The final dataset consists of 172 tender documents.

Each tender may include multiple lots, with each lot representing a specific procurement unit associated with a given MD or group of devices. In some cases, tenders include more than 100 lots; considering that multiple suppliers may participate in each lot, a single tender document can generate a large number of observations. As a result, it is possible to derive up to 200 - 300 distinct observations from a single document.

The preliminary analysis indicates that the available documentation appears sufficiently rich to support the development of a structured dataset. At this stage, the available evidence supports the feasibility of proceeding with the data extraction process.

Once the dataset was defined, a preliminary examination of the documents highlighted a significant degree of heterogeneity in structure, terminology, and data presentation formats. This variability represented a critical methodological challenge, as it limited the possibility of directly applying a uniform data extraction approach.

To address this issue, the documents were first classified into structurally homogeneous groups (Figure 1). This phase enabled the identification of recurring patterns in how evaluation data were organised and constituted the basis for a systematic extraction strategy. Three main document structures were identified: (i) documents in which each criterion is presented in a separate table containing the evaluations of all participants; (ii) documents where each participant is associated with an individual table reporting all criteria; and (iii) documents structured as a matrix, where rows correspond to criteria and columns to participants.

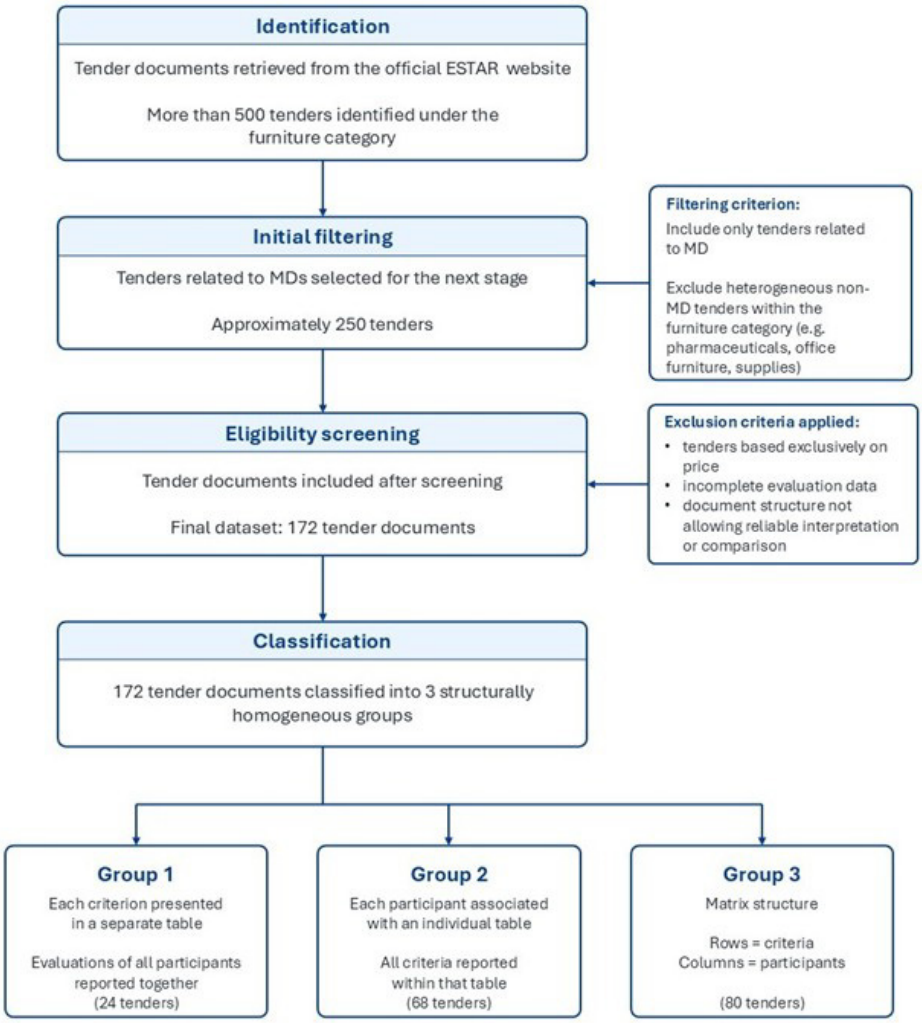


Figure 1: PRISMA diagram of the tender selection process

Based on this classification, tailored extraction schemas were defined for each document group, reflecting the specific organisation of information. The extraction process focused primarily on evaluation criteria, while also capturing complementary variables such as tender identifier, lot number, lot description, product category, participating suppliers, awarded supplier, scores, and final outcomes. This approach allowed the transformation of heterogeneous and partially unstructured documents into a consistent and analysable dataset.

The data extraction process is inherently iterative. Initial extraction is performed following a defined approach, but subsequent analysis of the extracted data may lead to refinements in the extraction strategy, requiring a return to earlier phases and adjustments to the procedure.

The outputs generated through AI-assisted extraction were systematically validated through manual comparison with the original tender documents, ensuring the accuracy and reliability of the collected data.

At the current stage, data have been extracted for a subset of documents corresponding to document structure (i). The same extraction procedure will be progressively adapted and extended to the remaining document groups in subsequent iterations.

3.2 Recoding of evaluation criteria into meta-criteria

Device-specific evaluation criteria were recoded into broader evaluative dimensions, referred to as meta-criteria. This phase addresses one of the main challenges in the analysis of MD procurement, namely the high level of specificity and heterogeneity of tender criteria, which limits comparability across different procurement procedures.

Although individual criteria often refer to highly specific technical or functional attributes, they nonetheless reflect more general evaluative logics. The coding process therefore enabled a shift from highly device-specific attributes to higher-level value dimensions that can be meaningfully compared across tenders.

Thus, meta-criteria were conceived as higher-level analytical categories capable of capturing recurring evaluative logics across different procurement contexts. Importantly, they do not replace the original criteria but provide a layer of abstraction. To preserve the original meaning assigned by the contracting authority, each criterion was retained in its original formulation and linked to its corresponding meta-criterion. This double-coding approach avoids information loss and ensures full traceability between the source document, the specific tender criterion, and the broader evaluative dimension identified during the analysis.

As noted, the research is still ongoing and currently situated in the second phase of the study, which focuses on the recoding of criteria into meta-criteria.

Accordingly, meta-criteria presented in Table 1 should be considered as preliminary results.

Table 1: List of identified meta-criteria

Meta-criterion	Description	Examples of criteria belonging to
Ergonomics, handling, and ease of use	It refers to the extent to which a MD is designed to be intuitive, comfortable, and practical for operators during use. This meta-criterion captures aspects related to physical interaction with the device – such as grip, manoeuvrability, and control – as well as the simplicity and efficiency of performing key tasks (e.g., positioning, insertion, or operation), with the aim of reducing user effort, improving procedural fluidity, and supporting safe and effective use in clinical settings.	· Ergonomics and ease of use of radiofrequency catheters · Ergonomics of the handle · Degree of ergonomics of the grip · Ease of one-handed application · Ease of use of the inflation system and maintenance of the achieved pressure · Ease of inserting the transfer catheter into the guide · Ease of positioning and possible removal · Greater practicality of use, understood as the presence and handling of the applicator
Versatility of use	It refers to the extent to which a MD meet different clinical needs and use scenarios, by offering a wide variety of configurations, sizes, functional options, formats, and compatible uses. This supports clinicians in selecting the most appropriate solution across diverse patients, procedures, and operational contexts.	· Breadth of the range of forceps offered (e.g., shape of the jaws, serration, presence or absence of fenestration) · Breadth of the range of jaw lengths · Breadth of the range of sizes and shapes · Breadth of the range of balloon lengths · Breadth of the range of lengths and diameters · Breadth of the field of view with the assembled kit · Possibility of use during radiotherapy sessions · Possibility of performing procedures at both lumbar and cervical levels

Meta-criterion	Description	Examples of criteria belonging to
Patient safety	It refers to the extent to which a MD minimizes risk associated with its use on patients, encompassing both the procedure itself and the post-procedure phase. It includes risks related to the behaviour of the healthcare operator or the patient (e.g., in the dosing and administration of substances used in combination with the MD) or linked to organizational aspects (e.g., storage methods of the technology, learning curve, operator training	· Safety features · Safety function (patient bypass function that can be activated manually and automatically) · Impermeability (protection from external agents) · Protective sheaths that facilitate passage through the mesh · Safety and perfect sealing of connections welded to the tube
Operator safety	It refers to the extent to which a MD minimizes risks of infection, radiation, etc., to which the healthcare operator is exposed during the MD use.	· Greater safety of closure after use · Availability of safety connectors with a leak-proof valve ("Clave Connector") for the administration of antineoplastic drugs · Safety system: ergonomic and easy to activate (the safety system features a tip protection designed to prevent accidental needle sticks, ensuring operator safety; it activates easily but is slightly bulky)
Interoperability	Ability of the device to connect, integrate, or function properly with other devices, instruments, equipment, accessories, or information systems already present in the clinical setting.	· Ease of connection to the bronchoscope · Ease of attachment to the aspirator · Possibility of direct connection to the ECG monitor without switch systems
Degree of clinical certainty	It refers to the extent to which the MD use is supported by available clinical evidence and scientific documentation. It captures the degree of certainty regarding the expected impact of the technology on patient health, considering the quality, quantity, and relevance of the existing evidence for the specific technology and target population. It may include, for example, the availability of clinical studies, follow-up data, peer-reviewed publications, evidence from comparable devices, and the robustness of the reported clinical outcomes.	· Clinical evaluation: presence of follow-up clinical studies and scientific literature, taking into account the high level of evidence · Literature: presence of clinical studies · Selected PubMed studies related to the offered product · Number of clinical studies published on PubMed related to the offered product · Presence of randomized clinical studies published in journals with an IF ≥ 10, demonstrating the possibility of discontinuing dual antiplatelet therapy after 30 days · Presence of randomized clinical studies on long-term outcomes, with reference to the number of patients enrolled in the studies and the IF of the journals in which they were published
Dimension and weight	It refers to the physical characteristics of a MD in terms of size, weight, volume, and spatial requirements. This meta-criterion captures the extent to which the device's dimensions are compatible with the clinical setting, procedural space, transport, storage, installation, or anatomical constraints.	· System size/footprint · Volume and weight · Anchoring and transport system; size and weight of the equipment · Prosthesis footprint in the aortic root or left ventricle · Balloon catheter size · Electrode size

Meta-criterion	Description	Examples of criteria belonging to
Material properties and biocompatibility	It refers to the MD physical, technical, biological, and construction-related characteristics. This meta-criterion captures the quality and suitability of the materials used, their technical properties, their compatibility with biological tissues, and the accuracy and reliability of the device's construction. It reflects how material composition and manufacturing quality contribute to the device's safety, durability, and suitability for its intended clinical use.	· Biocompatibility: the biocompatibility of the proposed devices will be evaluated · Materials used for the products: quality and technical characteristics of the materials intended to be used · Quality of the material

4 Discussion and Conclusion

This paper reports the empirical findings from phases 1 and 2 of a research project investigating how AI can support decision-making in MD procurement, with particular attention to its role in structuring unstructured tender data, identifying higher-level evaluative dimensions (meta-criteria), and enabling comparisons between AI-generated outputs and human procurement decisions. Specifically, phase 1 focused on the analysis and structuring of tender documents, while phase 2 involved the recoding of device-specific criteria into higher-level meta-criteria. Data extracted from the first documents analysed enabled the identification of an initial set of meta-criteria, including versatility of use, patient safety, operator safety, and interoperability. The development of AI-based value configurations and their systematic comparison with human procurement decisions is left to future work. These preliminary results provide the foundational knowledge base and evaluative dimensions required for subsequent AI-based modelling and decision-support analyses.

Once the overall methodology has been fully implemented, the study may contribute both theoretically and practically. From a theoretical perspective, it may provide insights into how AI can influence decision-making in the healthcare sector (Weissman, 2025; Cresswell et al., 2023), particularly in MD procurement. From a managerial perspective, several contributions can be envisaged. By making explicit the evaluative logics implicitly emerging from past procurement decisions, AI-based value configurations could provide decision-makers with a structured representation of how decisional criteria are combined in practice, beyond those formally stated in tender documents (Siciliani et al., 2023). Moreover, comparing human decisions with these configurations could support the assessment of decision consistency and robustness, highlighting potential misalignments that may warrant managerial attention rather than necessarily indicating suboptimal choices (Guida et al., 2023; Siciliani et al., 2023). Such comparisons

5 Future Directions

As noted, this study is part of ongoing research. At present, the work is positioned within the second phase of the methodology, namely the identification of meta-criteria. However, the first phase is also still in progress, as data extraction has so far been completed for only one of the identified groups of documents. It is reasonable to expect that the extraction of data from the remaining document groups may reveal additional features, potentially leading to a refinement of subsequent outputs – particularly those related to Phase 2 – in order to achieve a more robust and comprehensive definition of meta-criteria.

Once Phase 2 is completed, the research will proceed to Phase 3, in which meta-criteria will be combined to construct AI-based value configurations within specific procurement contexts. Phase 4 will then compare these configurations with observed human procurement decisions, relying primarily on predictive performance metrics and on the systematic analysis of cases in which AI-based recommendations and human decisions diverge. This comparison will make it possible to assess alignment, identify divergences, and support a more reflexive evaluation of procurement decision-making processes.

More specifically, the next stages will not only focus on the comparison between AI-generated outputs and human decisions, but also on the critical phase of transforming unstructured tender documentation into a structured and analysable knowledge base. This process is expected to contribute to improving comparability, transparency, and consistency in procurement decisions, while also providing insights into the underlying evaluative logics that shape them.

6 Ethics declaration

Ethical clearance was not required for this study, as the research was based exclusively on publicly available tender documents.

7 AI Declaration

AI tools were used in two limited and controlled ways within this study. First, AI-based tools were employed for language editing to improve clarity and readability of the manuscript. Second, as described in the methodology, data extraction from tender documents was performed using a generative AI chatbot based on a family of large language models. This process was conducted under strict human supervision, and all extracted data were systematically validated through manual comparison with the original tender documents to ensure accuracy and reliability. No AI system was used for autonomous analysis without human oversight.

References

Aboelazm, K., & Dganni, K. (2025). Public procurement contracts futurity: Using of artificial intelligence in a tender process. Corporate Law & Governance Review. https://doi.org/10.22495/clgrv7i1p6

Acikalin, U. U., Gorgun, M. K., Kutlu, M., & Tas, B. K. O. (2024). How you describe procurement calls matters: Predicting outcome of public procurement using call descriptions. *Natural Language Engineering*, *30*(6), 1255–1276. https://doi.org/10.1017/S135132492300030X

Aljafari, A. (2021). Quality and Efficiency of Health Care Systems During the COVID-19 Pandemic: Assessment of G20 Unified Procurement Systems for Medical Equipment. https://doi.org/10.21203/rs.3.rs-844381/v1

Allal-Chérif, O., Simón-Moya, V., & Cuenca Ballester, A. C. (2021). Intelligent purchasing: How artificial intelligence can redefine the purchasing function. *Journal of Business Research*, *124*, 69–76. https://doi.org/10.1016/j.jbusres.2020.11.050

Atella, V., & Decarolis, F. (2019). *Procuring medical devices: Evidence from Italian public tenders* (CEIS Working Paper No. 472). SSRN. https://doi.org/10.2139/ssrn.3467466

Balkan, D., & Akyuz, G. (2025). Artificial intelligence (AI) and machine learning (ML) in procurement and purchasing decision-support (DS): a taxonomic literature review and research opportunities. *Artificial Intelligence Review*, *58*. https://doi.org/10.1007/s10462-025-11336-1

Borsoi, L., Callea, G., Amatucci, F., Marsilio, M., & Tarricone, R. (2017). Centralized Procurement of Medical Devices In Italy: A Methodology Towards Standardized Tender Dossier. Value in Health, 20. https://doi.org/10.1016/j.jval.2017.08.1110

Bosmans, H., Zanca, F., & Gelaude, F. (2021). Procurement, commissioning and QA of AI based solutions: An MPE's perspective on introducing AI in clinical practice. *Physica Medica: European Journal of Medical Physics*, *83*, 257–263. https://doi.org/10.1016/j.ejmp.2021.04.006

Cangelosi, M., Chahar, A., & Eggington, S. (2023). Evolving use of health technology assessment in medical device procurement—Global systematic review: An ISPOR Special Interest Group report. *Value in Health*, *26*(11), 1581–1589. https://doi.org/10.1016/j.jval.2023.06.005

Cernăzanu-Glăvan, C., & Bulzan, A. (2025). Explainable AI and Fuzzy Linguistic Interpretation for Enhanced Transparency in Public Procurement: Analyzing EU Tender Awards. Mathematics. https://doi.org/10.3390/math13132215

Cresswell, K., Rigby, M., Magrabi, F., Scott, P., Brender, J., Craven, C. K., Wong, Z. S.-Y., Kukhareva, P., Ammenwerth, E., Georgiou, A., Medlock, S., de Keizer, N. F., Nykänen, P., Prgomet, M., & Williams, R. (2023). The need to strengthen the evaluation of the impact of artificial intelligence-based decision support systems on healthcare provision. *Health Policy*, *136*, 104889. https://doi.org/10.1016/j.healthpol.2023.104889

Diaconu, K., Chen, Y.-F., Cummins, C., Jimenez Moyao, G., Manaseki-Holland, S., & Lilford, R. (2017). Methods for medical device and equipment procurement and prioritization within low- and middle-income countries: Findings of a systematic literature review. *Globalization and Health*, *13*, Article 59. https://doi.org/10.1186/s12992-017-0280-2

Guida, M., Caniato, F., Moretto, A., & Ronchi, S. (2023). The role of artificial intelligence in the procurement process: State of the art and research agenda. *Journal of Purchasing and Supply Management*, *29*(2), Article 100823. https://doi.org/10.1016/j.pursup.2023.100823

Hinrichs-Krapels, S., Ditewig, B., Boulding, H., Chalkidou, A., Erskine, J., & Shokraneh, F. (2022). Purchasing high-cost medical devices and equipment in hospitals: a systematic review. BMJ Open, 12. https://doi.org/10.1136/bmjopen-2021-057516

Khan, S., Hoodbhoy, Z., Raja, M., Kim, J., Hogg, H., Manji, A., Gulamali, F., Hasan, A., Shaikh, A., Tajuddin, S., Khan, N., Patel, M., Balu, S., Samad, Z., & Sendak, M. (2024). Frameworks for procurement, integration, monitoring, and evaluation of artificial intelligence tools in clinical settings: A systematic review. *PLOS Digital Health*, *3*. https://doi.org/10.1371/journal.pdig.0000514

Kufile, O. T., Akinrinoye, O. V., Umezurike, S. A., Ejike, O. G., Otokiti, B. O., & Onifade, A. Y. (2022). Advances in data-driven decision-making for contract negotiation and supplier selection. International Journal of Multidisciplinary Research and Growth Evaluation, 3(2), 831-842 https://doi.org/10.54660/.IJMRGE.2022.3.2.831-842

Long, P., Lu, L., Chen, Q., Chen, Y., Li, C., & Luo, X. (2023). Intelligent selection of healthcare supply chain mode: An Applied research based on artificial intelligence. *Frontiers in Public Health*, *11*, Article 1310016. https://doi.org/10.3389/fpubh.2023.1310016

Miller, F. A., Lehoux, P., Peacock, S., Rac, V. E., Neukomm, J., Barg, C., Bytautas, J. P., & Krahn, M. (2019). How procurement judges the value of medical technologies: A review of healthcare tenders. *International Journal of Technology Assessment in Health Care*, *35*(1), 50–55. https://doi.org/10.1017/S0266462318003756

Muehlematter, U. J., Daniore, P., & Vokinger, K. N. (2021). Approval of artificial intelligence and machine learning-based medical devices in the USA and Europe (2015–20): A comparative analysis. *The Lancet Digital Health*, *3*(3), e195–e203. https://doi.org/10.1016/S2589-7500(20)30292-2

Nai, R., Meo, R., Morina, G., & Pasteris, P. (2023). Public tenders, complaints, machine learning and recommender systems: A case study in public administration. *Computer Law & Security Review*, *51*, Article 105887. https://doi.org/10.1016/j.clsr.2023.105887

Rainy, T. A., & Chowdhury, A. R. (2022). The role of artificial intelligence in vendor performance evaluation within digital retail supply chains: A review of strategic decision-making models. *American Journal of Scholarly Research and Innovation*, *1*(1), 220–248. https://doi.org/10.63125/96jj3j86

Siciliani, L., Taccardi, V., Basile, P., Ciano, M., & Lops, P. (2023). AI-based decision support system for public procurement. Inf. Syst., 119, 102284. https://doi.org/10.1016/j.is.2023.102284

Tallarico, S., Aloini, D., Dulmin, R., Lazzini, S., Mininno, V., & Pellegrini, L. (2022). Health technology assessment of medical devices: Overcoming the critical issues of current assessment. Journal of Multi-Criteria Decision Analysis, 29(1–2), 150–172. https://doi.org/10.1002/mcda.1764

Angehrn, Z., Haldna, L., Zandvliet, A. S., Gil Berglund, E., Zeeuw, J., Amzal, B., Cheung, S. Y. A., Polasek, T. M., Pfister, M., Kerbusch, T., & Heckman, N. M. (2020). Artificial intelligence and machine learning applied at the point of care. *Frontiers in Pharmacology*, *11*, Article 759. https://doi.org/10.3389/fphar.2020.00759

Weissman, G. E. (2025). Evaluation and regulation of artificial intelligence medical devices for clinical decision support. *Annual Review of Biomedical Data Science*, *8*(1), 81–99. https://doi.org/10.1146/annurev-biodatasci-103123-095824

Zorina, Y., & Bikir, A. (2025). Automation vs. intelligence: Limits and opportunities of AI in e-procurement platforms. *The American Journal of Management and Economics Innovations*, *7*(8), 145–154. https://doi.org/10.37547/tajmei/Volume07Issue08-13

Training Project Managers through Games and Simulations: A Literature Review

Teresa Anna Rita Gentile[1], Ernesto De Nito[2]
[1] University of Padova, Padova, Italy
[2] University of Salerno, Italy

Abstract

The present research aims to verify how teaching methodologies based on games and simulations promote the training of soft skills of project managers. Specifically, through a review of the literature, the types of soft skills that project managers can acquire if they follow training courses based on simulations and games are identified.
The methodological approach we have adopted is quantitative and marked by two phases. In the first phase, we identify the topics consistent with the objective of the survey, with the help of the Scopus database, for the period from 1999 to 2024. In the second phase, we analyze, firstly, the content of the articles consistent with reference to teaching methodologies based on games and simulations for the training of project managers and, secondly, we try to identify what types of soft skills they develop.
As regards the main descriptive data of the first phase, two main elements emerge. First of all, with reference to the overall figure of publications found on the subject (n = 298) for each year, it emerges that starting from the period 2015 there is an increasing increase in publications and in 2023 there are as many as n = 66 publications. Secondly, most of the documents found are present in the area of Social Sciences (n=122), Business, Management and Accounting (n=99).
With reference to the second phase, it emerges that many methods based on simulations and business games allow project managers to acquire specific soft skills such as leadership, communication skills, teamwork, to be able to develop the decision-making process and to be able to solve problems.
This article has theoretical implications, as it aims to fill the gap in the current literature on simulations and games to encourage the study of soft skills in project management.
In addition, it has important implications from a practical point of view, as it allows HR managers to plan training courses characterized by simulations and games that include the development of soft and hard skills in an integrated way.

Keywords: Project manager, training, soft skills, games, simulations

Paper type: Academic Research Paper

1 Introduction

Knowledge-intensive organizations (KIOs) are characterized by being places where people possess the skilled skills and experiences necessary to perform complex tasks that require autonomy, judgment, and the ability to solve complex tasks (Makami and Marche, 2016). Furthermore, in these KIOs there is an increasing tendency to work with "project-based" methods because considerable amounts of work are managed that require the presence of the figure of the project manager (Cheng, et al., 2005). As we know, the project manager should possess several skills to be able to lead the group he is responsible for in the successful achievement of the project objectives (Project Management Institute, 2017; Stevenson and Starkweather, 2010). Precisely for this reason, the project manager should be equipped with a combination of skills that include technical but also transversal skills such as: experience, communication skills, adequate training, knowledge of project management methodologies and the functioning of certain business processes, organizational skills, problem solving skills, time management and team management skills and decision-making skills (Pons et al., 2024; Wyskwarski, 2021). At the same time, according to Alvarenga et al. (2020), soft skills such as communication skills, commitment, and leadership would be very important for the success of the project and would be more important than technical skills. Thanks to these soft skills, the project manager would be able to better interact with the other members of the group to achieve the behavior of the latter in accordance with the project objectives (Xue, et al., 2020). According to this approach, it is appreciable not only to recognize soft skills, but also to identify which teaching methodologies are easier to use to train and develop these soft skills in project managers.

In the literature it is discussed whether, in order to consistently train the manager in the diligent execution of the project, it is necessary to opt for training programs implemented in the classroom (Buganza et al. 2013; Siddique et al. 2022) or for those carried out in workplaces where games and simulations are used (Long and Mawdesley, 2010). Many studies have looked at training courses for PMs delivered through gamification. For example, some articles have focused on how game-based techniques allow future PMs to acquire these skills

(Barnabè et al. 2023; Baccarani et al. 2014) or to allow project managers already in service to develop their skills (Abidin and Fathi, 2019). However, the analysis of the main literature shows that the focus on games and simulations for the acquisition of the project manager's soft skills is a little-covered topic (Kearney, et al., 2024).

Precisely for this reason, the aim of this article is to verify, through a review of the literature, whether teaching methodologies based on games and simulations support the acquisition of soft skills by project managers.

The article is divided like this. First of all, the main reference literature is reported, then the paragraph on methodology is exposed, then the main results obtained are shown and discussed and, finally, the conclusions are presented.

2 Main Literature

When we talk about training project managers in soft skills, such as those related to leadership, communication, conflict management and teamwork, the development of critical thinking or the ability to make decisions, we refer to training programs that can take place in the classroom or on the job. In the first case, the focus is on the acquisition of knowledge (Al-Nabae and Sammani, 2019), as students are allowed to learn a lot in theoretical terms, also through self-study, using materials and suggestions provided by the teacher (Johar et al. 2016). In on-the-job learning, on the other hand, there is a tendency to make project managers understand how to correctly carry out certain activities, especially if they are complex projects (Thamhain, 1991), as they are placed in real work situations, through simulations and games (Medina, 2006).

This modality allows project managers to learn quickly and to develop networking and dialogue with other colleagues who are involved in projects (Palotie et al., 2017). In fact, the use of games and simulations can be effective in making a project manager understand how to join a team, create a database or set up a new process to relate to other professional colleagues (de los Ríos Carmenado et al. 2012), acquire both technical and soft skills (Hellström et al., 2023; Seager et al., 2011) and meet specific learning or training needs (de Freitas, 2006).

Specifically, games are teaching methods that allow to re-propose real work situations in which players following a strategy and through collaboration with other players are able to make a decision (Gooding and Zimmerer, 1980; Piraveenan, 2019). In concrete terms, it happens that project managers in training put themselves in the shoes of players, formulate some strategies, then on the basis of the assignment of some values, they are able to collect information, analyze it and make decisions (Potter, 1968) and also have the opportunity to experience the consequences of these choices on the other actors involved (Martin et al. 2014).

In this way, project managers participating in the games can put their theoretical knowledge into practice and refine their soft skills and be able to develop specific skills that are difficult to acquire with traditional teaching methods (Bočková et al. 2015).

Instead, simulation-based teaching methods place aspiring project managers in a situation similar to the one they might encounter in the workplace (Noe et al. 2006), so that they can experience the activities of managing a project (Nascimento et al., 2013) and solve complex situations (Castronovo et al. 2017).

These forms of learning are characterized by the use of special simulators which, also with the help of suitable software (Martin et al. 2014), are able to generate unstructured scenarios and learning environments suitable for guaranteeing useful feedback to solve the problems that project managers may encounter in the work context (Akiyoshi et al. 2012). In concrete terms, project managers are assigned a simulation of a project to be implemented in a group, also through the use of platforms, in which they are asked to solve some problems and achieve the goal of the project through the exploitation of their hard and soft skills (Jaccard and Riboni, 2010).

Through this simulation-based method, the project manager can develop decision-making and critical thinking techniques and can acquire greater teamwork and problem-solving skills (Chernikova et al. 2020).

Yet, the topic of teaching methodologies characterized by simulations and games to make project managers acquire transversal skills is little covered in the literature (Long and Mawdesley, 2010; Kearney et al., 2024).

In light of this, the aim of this research is to verify, through a review of the main literature, whether teaching methodologies based on games and simulations can facilitate the training of soft skills of project managers.

3 Methodology

We used a quantitative methodological approach characterized by two phases.

In the first step, we identified the items through the use of the Scopus database. We used this database over others because Scopus in its research considers many sources, includes open access publications and publications written in languages other than English and also proposes a longer time frame (Valderrama-Zuriàn et al. 2015).

We queried the Scopus database by entering the following search string: ("project manag") AND ("training") OR ("education") AND ("soft skill") OR ("behavioral") AND ("game") AND ("simulation"), indicating as the period 1999-2024. We chose this long time frame for our investigation because there has been an increased use of the Project Manager figure in organizations around the world. We have considered for the research only articles, conference proceedings and book chapters written in English. In addition, we excluded those articles that contained keywords far from the string of our search.

On the basis of the identified articles (n=1,770) we built the Prisma model (Liberati et al., 2009; Kelly et al., 2016) according to the scheme better illustrated in figure 1.

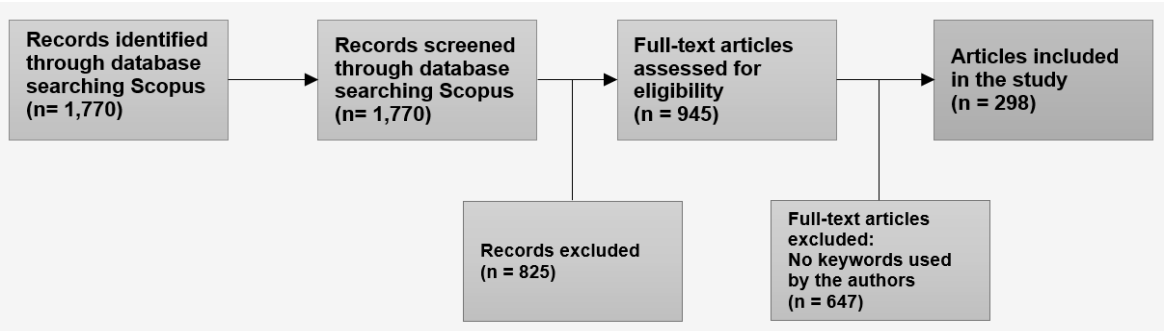


Figure 1: PRISMA flowchart.
Sources: adapted from Liberati et al., 2009; Kelly et al., 2016

In the second phase we studied, within the final sample of the articles found (n=298), the articles dealing with two specific topics: (i) Social Sciences (n=122) and (ii) Economics, Management and Accounting (n=99), as their contents are more consistent with the objectives of the research. We proceeded first with the study of the abstracts identified in these two topics (n=211) to verify their consistency with the research objective. Thus, we have deepened the content of the coherent articles identified within the aforementioned two themes, considering two fundamental aspects. Firstly, we select, from the relevant contributions, those that use methodologies based on games and simulations in the training of project managers and, secondly, we try to identify what types of soft skills they develop.

4 Initial Results and Discussion

With reference to the first phase of the research, from the data collected by the Scopus database and reported in the PRISMA flowchart (figure 1) it emerges that the n=1770 articles found in the aforementioned database became n=1395 because only those written in English were considered, then selecting only articles, conference proceedings and book chapters we arrived at n=945 items. Finally, excluding articles that had keywords other than those indicated in our search string, we identified a total sample of n= 298 documents.

Within this sample, the following descriptive data emerged.

First, Figure 2 shows that the number of articles published in journals (n=220) is higher than those of book chapters (n=58) found and those referring to conference proceedings (n=20).

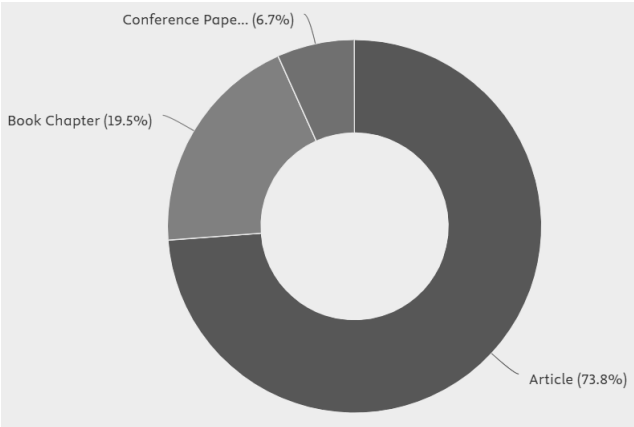


Figure 2: Number of documents found by type 1999-2024.
Source: Scopus Elsevier B.V. (2026)

Furthermore, secondarily, it emerges from figure 3 that most of the documents found are present in the areas of Social Sciences (n=122), Business, Management and Accounting (n=99), Computer Science (n=89) and Engineering (n=70). In the other areas there is a lower number of documents found.

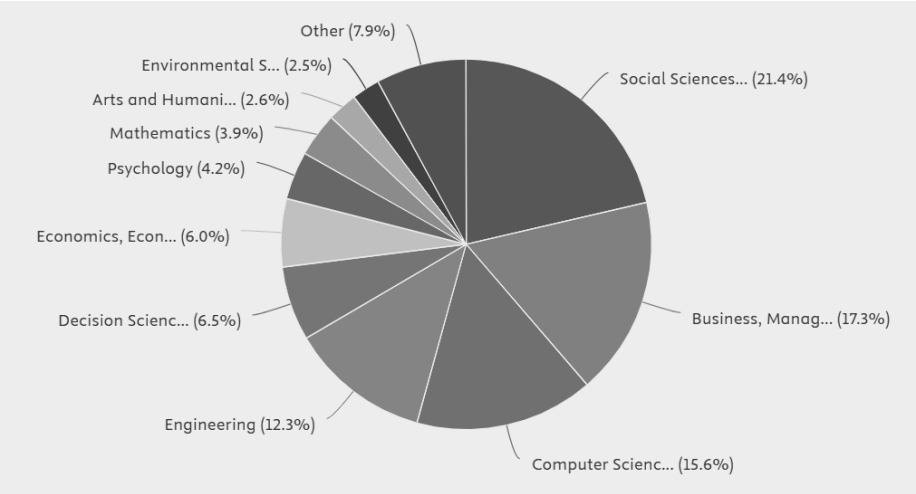


Figure 3: Number of documents by subject area 1999-2024.
Source: Scopus Elsevier B.V. (2026)

Thirdly, with reference to the overall data of publications on the subject for each year in the period 1999-2024 (see figure 4), it can be observed that at the beginning of the period in 2000 there was a feeble attention to the phenomenon (n=5 items) which almost disappeared in the following years until it had a slight decrease in 2010 (n=3 documents). Only starting from 2014 (n=6 articles) and 2015 (n=9 items) is there an increasing increase in publications and, in fact, in 2023 there are as many as n=66 publications.

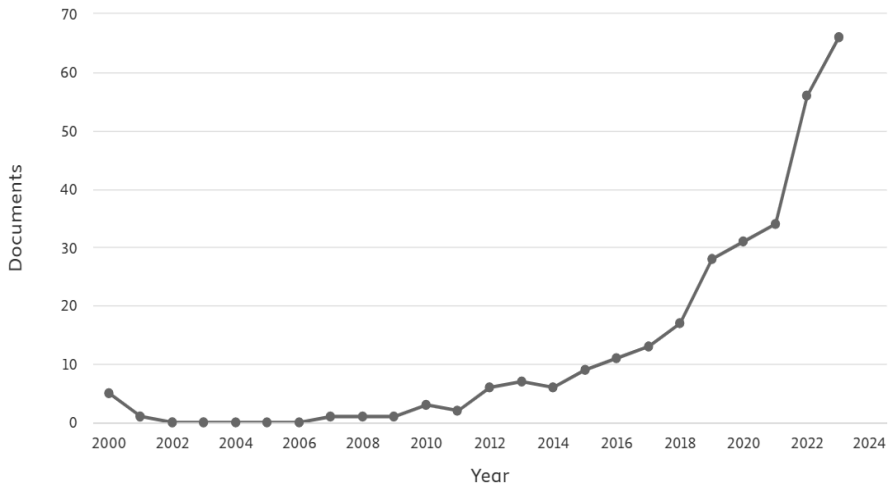


Figure 4: Publications for each year in the research 1999-2024.
Source: Scopus Elsevier B.V. (2026)

As regards the countries of origin of the authors or publishers - in the case of books or conference proceedings - (figure 5) it can be deduced from this analysis that most of the publications collected are related to the United States (n=60), followed by China (n=35) and the United Kingdom (n=27). For Europe we have 10 items that belong equally to Germany and Spain, while only n=8 items belong to France and n=7 items to Italy. These last two data are not present in the aforementioned graph.

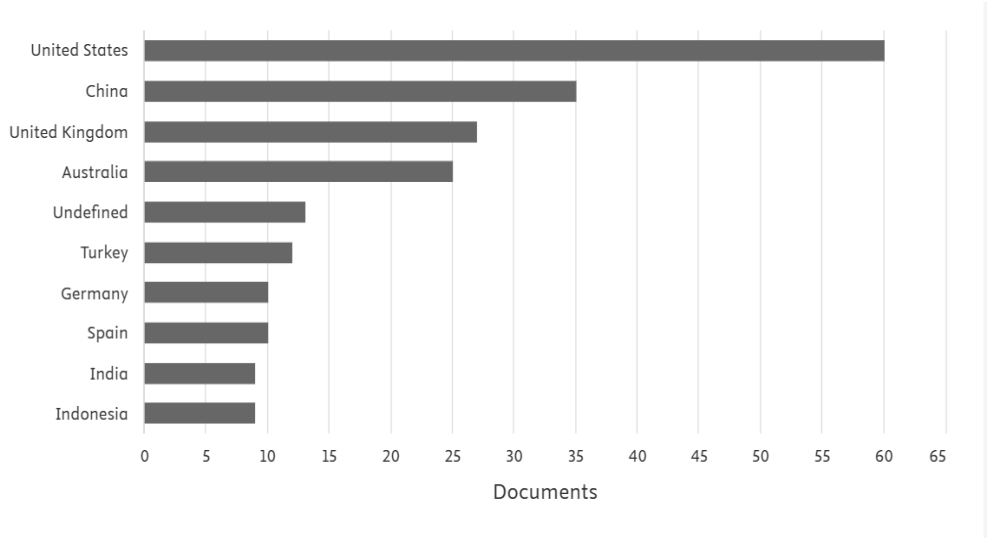


Figure 5: Number of documents by author/publisher membership 1999-2024.
Source: Scopus Elsevier B.V. (2026)

In the second phase of our research, we analyzed the articles (n=221) related to the topics: (i) Social sciences and (ii) Economics, management and accounting which are the most present in the field of the articles found according to figure 3.

We have carried out a meticulous analysis of the abstracts of the above-mentioned articles, from which it emerged that only n=19 articles are consistent with our investigation, as they consider simulations and games as the main methodologies for the training of PMs, as can be seen from table 1 below

Table 1: Number of articles on simulations and business games found.

No.	Author(s)	Year	Title	Game used	Type of soft skill developed
1	Akiyoshi et al.	2012	*A project manager skill-up simulator: towards problem solving-based learning*	*Simulation game*	*Problem-solving*
2	Barnabè et al.	2023	*Critical thinking skills enhancement through system dynamics-based games: insights from the project management board game project*	*Simulation game*	*Critical thinking*
3	Bonazzi et al.	2013	*Analysis of serious games implementation for project management courses*	*Serious game*	*Problem-solving*
4	Castronovo et al.	2017	*Developing problem-solving skills in construction education with the virtual construction simulator*	*Simulation game*	*Problem-solving*
5	Alves	2022	*The silo game: a simulation on interdisciplinary collaboration*	*Simulation game*	*Team working*
6	Geithner and Menzel	2016	*Effectiveness of learning through experience and reflection in a project management simulation*	*Simulation game*	*Team working*
7	Gonen and Israeli	2016	*New results using project management simulation game*	*Simulation game*	*Decision making*
8	Habouba-Belinky and Parush	2020	*Simulator-triggered debriefing facilitates team cognition and performance in simulator-based training*	*Simulation game*	*Team working*
9	Hamada et al.	2013	*Reference operation generation method on project manager skill-up simulator*	*Simulation game*	*Problem-solving*
10	Jaccard et al.	2022	*How might serious games trigger a transformation in project management education: lessons learned from 10 years of experimentations*	*Serious game*	*Communication Decision making*
11	Jaccard and Riboni	2010	*Sim ProJet: an innovative simulation platform for experiential learning in project management*	*Simulation game*	*Team working*

No.	Author(s)	Year	Title	Game used	Type of soft skill developed
12	Korman and Johnston	2010	*Enhancing construction management education through the use of a virtual construction company simulation system*	*Simulation game*	*Team working*
13	Monga et al.	2017	*Context-rich digital games for better learnability in the IT project management context*	*Simulation game*	*Decision making*
14	Nascimento et al.	2013	*Project management simulation portal: proposal, features and construction process*	*Simulation game*	*Decision making Leadership*
15	Rumeser and Emsley	2019	*Can serious games improve project management decision making under complexity?*	*Simulation game*	*Decision making*
16	Sluka	2019	*Project managers competences game: how does it work*	*Simulation game*	*Team working*
17	Zoroja et al.	2019	*Developing leadership in a simulation games environment*	*Simulation game*	*Leadership*
18	Zwikael and Gonen	2007	*Project execution game (PEG): training towards managing unexpected events*	*Simulation game*	*Decision making Problem-solving*

In the light of a brief reading of these 19 articles, the following reflections emerge.

First of all, it would seem to be the possibility of making project managers, especially those who work in certain sectors, acquire soft skills related to problem solving and the ability to make decisions through games and simulations (Akiyoshi et al., 2012; Bonazzi et al., 2013; Castronovo et al., 2017; Gonen and Israeli, 2016; Hamada et al., 2013; Monga et al., 2017; Nascimento et al., 2013; Rumeser and Emsley, 2019; Zwikael and Gonen, 2007). This need arises from the fact that a good PM can be faced with problems of various kinds during the implementation of projects and only if he possesses these "soft skills" can he achieve the goal of the project smoothly (Crawford, 2000).

Secondly, many simulation- and game-based methods allow project managers to acquire soft skills related to leadership and the promotion of critical thinking (Barnabè et al., 2023; Zoroja et al., 2019).

In addition, many games and simulations allow project managers to acquire new techniques to be able to work in a team (Alves, 2022; Habouba-Belinky and Parush, 2020; Geithner and Menzel, 2016; Korman and Johnston, 2010; Sluka, 2019) or to develop and improve communication skills (Jaccard et al., 2022).

Finally, it can be said that training techniques based on games and simulations, through the guaranteed safe environment, offer the possibility for aspiring project managers to gain practical experience, to be able to proceed by trial and error, to work in teams and to take an active role (Fahmi et al., 2021) and to improve the learning and development of project management (Abidin and Fathi, 2019).

5 Conclusions

This article aims to detect, through a review of the literature, how teaching methodologies based on games and simulations are able to support the training of soft skills of project managers.

The main results obtained show that in the period studied, from 1999 to 2024, simulation-based methods and business games allowed aspiring PMs to develop particular soft skills, such as those related to leadership and critical thinking (Zoroja et al., 2019); or those of problem-solving and decision-making (Bonazzi et al., 2013) or, finally, those related to teamwork (Sluka, 2019) or communication (Jaccard et al., 2022).

In any case, we have been able to see that methods based on play and simulations, although they have the advantage of allowing aspiring PMs to proceed in a safe environment, having the possibility of making trials and mistakes, require a lot of time for their conception and for the construction of the phases of the game (Fahmi et al., 2021). In addition, these teaching methods could generate forms of competition with the other members of the group among the different PMs involved in the game, since, due to the identification with the characters to be played in the game, the desire to win instinctively emerges in each of them (Farooq et al., 2022).

This article has theoretical implications, as it aims to fill the gap in the current literature on simulations and games to encourage the study of soft skills in project management. In addition, it could try to link the theme of the project manager's human resources development within the Kolb (1984) learning cycle, trying to re-propose the four phases and identify the new knowledge learned.

In addition, it has important implications from a practical point of view, as it could help HR managers to plan training courses characterized by teaching methods based on the use of simulations and games to develop not only soft skills, but also to think about designing courses for the integrated development of soft and hard skills.

The main limitation of this study is to be linked not only to the fact that it presented data without statistics with specific formulas, but also to the type of database used for the review of the literature. In fact, Scopus, although it enjoys excellent notoriety and reliability, has some critical issues in data generation, as it does not consider resources that have not yet been indexed.

Finally, we decided to consider only publications in English as a survey sample, ignoring other languages, since we assumed that the theme of the project manager was more felt by Western culture.

There is the intention for the near future to implement an empirical analysis involving some PMs, in the construction sector, to understand what types of soft skills need to be increased and to understand whether or not the use of teaching methods based on simulations and games is suitable for achieving this goal.

6 Ethics Declaration

The authors did not use ethical data that required the approval of the relevant committee.

7 AI Declaration

The authors did not use artificial intelligence tools to develop the article's content.

References

Abidin, N. A. Z., and Fathi, M. S. (2019) "The potential use of game-based virtual reality training for construction project managers", *Malaysian Construction Research Journal*, Vol. 76.

Akiyoshi, M., Samejima, M., and Komoda, N. (2012) "A Project Manager Skill-up Simulator Towards Problem Solving-based Learning." *KMIS*.

Al-Nabae, M. G., and Sammani D. (2019) "Training methods for the development of project team members", *International Journal of Innovation and Industrial Revolution*, Vol. 1, No. 1, pp. 01-12.

Alvarenga, J. C., Branco, R. R., Guedes, A. L. A., Soares, C. A. P. and Silva, W. D. S. E. (2020) "The project manager's key competencies for project success", *International Journal of Project Management in Business*, Vol. 13, No. 2, pp 277-292.

Alves, T. (2022) "The Silo Game: a simulation on interdisciplinary collaboration." Annual Conference Of The International Group For Lean Construction (Vol. 30).

Baccarani, C., Bayart, C., Bertezene, S., Vallat, D. and Martin, J. (2014) "Serious games: leverage for knowledge management", *The TQM Journal*, Vol. 26, No. 3, pp 235-252.

Barnabè, F., Armenia, S., Nazir, S. and Pompei, A. (2023) "Critical thinking skills enhancement through system dynamics-based games: Insights from the project management board game project", *Systems*, Vol. 11, No. 11, pp 554.

Bočková, K. H., Sláviková, G. and Gabrhel, J. (2015) "Game theory as a tool of project management" *Procedia-Social and Behavioral Sciences*, Vol. 213, pp. 709-715.

Bonazzi, R., Missonier, S., Jaccard, D., Bienz, P., Fritscher, B. and Fernandes, E. (2013). "Analysis of serious games implementation for project management courses" *Information Systems: Crossroads for Organization, Management, Accounting and Engineering: ItAIS: The Italian Association for Information Systems*.

Buganza, T., M. Kalchschmidt, E. Bartezzaghi, and Amabile, D. (2013) "Measuring the impact of a major project management educational program: The PMP case in Finmeccanica", *International Journal of Project Management*, Vol. 31, No. 2, pp. 285-298.

Castronovo, F., Van Meter, P., Zappe, S., Leicht, R., and Messner, J. (2017), "Developing problem-solving skills in construction education with the virtual construction simulator", *International journal of engineering education*, Vol. 33, No. 2, pp 831-846.

Cheng, M. I., Dainty, A. R. and Moore D. R. (2005) "What Makes a Good Project Manager?", *Human Resource Management Magazine*, Vol. 15, No. 1, pp 25-37.

Chernikova, O., Heitzmann, N., Stadler, M., Holzberger, D., Seidel, T., and Fischer, F. (2020). "Simulation-based learning in higher education: A meta-analysis", *Review of educational research*, Vol. 90, No. 4, pp 499-541.

Crawford, L. (2000) "Project management competence for the new millenium", *Proceedings of 15th World Congress on Project Management.*

de Freitas S. I. (2006) "Using games and simulations for supporting learning", *Learning, Media and Technology*, Vol. 31, No. 4, pp 343-358.

de los Ríos Carmenado, I., Rodríguez, B. F. and Gajardo F. G (2012) "Methodological proposal for teamwork evaluation in the field of project management training", *Procedia-Social and Behavioral Sciences*, Vol. 46, pp. 1664-1672.

Fahmi Hassan, A., Haghighi-Rad, F., and Abtahi, A. R. (2021) "Enabling construction project managers through a management game." *Industrial and Commercial Training*, Vol. 53, No. 4, pp 343-365.

Farooq, M. S., Hamid, A., Alvi, A., and Omer, U. (2022) "Blended learning models, curricula and gamification in project management training", *Access IEEE*, Vol. 10, pp. 60341-60361.

Geithner, S., and Menzel D. (2016) "Effectiveness of learning through experience and reflection in a project management simulation", *Simulation and Games*, Vol. 47, No. 2, pp 228-256.

Gonen, A., and Israeli, U. (2016) "New results using project management simulation game." *2016 IEEE International Conference on Management of Innovation and Technology*. IEEE.

Gooding, C., and Zimmerer, T. W. (1980) "The use of specific industry gaming in the selection, orientation and training of managers", *Human Resource Management*, Vol. 19, No. 1, pp. 19.

Habouba-Belinky, I., and Parush, A. (2020) "Simulator-triggered debriefing facilitates team cognition and performance in simulator-based training", *Journal of Interactive Learning Research*, Vol. 31, No. 1, pp 29-48.

Hamada, K., Akiyoshi, M., and Samejima, M. (2013) "Reference Operation Generation Method on Project Manager Skill-up Simulator" *International Joint Conference on Knowledge Discovery, Knowledge Engineering, and Knowledge Management*.

Hellström, M. M., Jaccard, D. and Bonnier, K. E. (2023) "A systematic review on the use of serious games in project management education", *International Journal of Serious Games*, 2.

Jaccard, D., and Riboni, S. (2010) "Sim projet: an innovative simulation platform for experiential learning in project management." *Proceedings of CSEDU 2010-proceedings of the Second International Conference on Computer Supported Education.*

Jaccard, D., Bonnier, K. E. and Hellström, M. (2022) "How might serious games trigger a transformation in project management education? Lessons learned from 10 Years of experimentations", *Project Leadership and Society*, Vol. 3, pp. 100047.

Johar, S., Nik Ibrahim, N. L., and Che Ani, A. I. (2016) "The assessment and application of student competency in land survey, building and measured drawing 'course", *Pertanika Journal of Social Sciences and Humanities*, Vol. 24, No. 4, pp 269-282.

Kearney, J., Bond-Barnard, T. and Chugh R. (2024) "Soft skills and learning methods for 21st-century project management: a review", *International Journal of Information Systems and Project Management*, Vol. 12, No. 4, pp 5-20.

Kelly, S. E., Moher, D., and Clifford, T. J. (2016) "Quality of conduct and reporting in rapid reviews: an exploration of compliance with PRISMA and AMSTAR guidelines", *Systematic reviews*, Vol. 5, No. 1, pp 79.

Korman, T., and Johnston, H. (2010) "Enhancing construction management education though the use of a virtual construction company simulation system." *The 2nd International Symposium on Engineering Education and Educational Technologies: EEET.*

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P., ... and Moher, D. (2009) "The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration", *Bmj*, 339.

Long, G., and Mawdesley, M. (2010) "Teaching construction management through simulation games: improving the learning experience", *EDULEARN10 Proceedings*. IATED.

Magano, J., Silva, C., Figueiredo, C., Vitória, A., Nogueira, T. and Pimenta Dinis, M. A. (2020) "Generation Z: Adapting Soft Project Management Skills: A Blended Method Approach", *Educational Sciences*, Vol. 10, No. 7, pp 187.

Makani, J., and Marche, S. (2010) "Towards a typology of knowledge-intensive organizations: determinant factors", *Knowledge Management Research & Practice*, Vol. 8, No. 3, pp 265-277.

Martin, B. O., Kolomitro, K., and Lam, T. C. (2014) "Training methods: A review and analysis", *Human Resource Development Review*, Vol. 13, No. 1, pp. 11-35.

Medina R. (2006) *Personnel & Human Resources Management*, Rex Book Store.

Medina, R., and Medina, A. (2017) "Managing competence and learning in knowledge-intensive, project-intensive organizations: A case study of a public organization", *International Journal of Managing Projects in Business*, Vol. 10, No. 3, pp 505-526.

Monga, C., Jain, J., Kumar, S. and Sandeep, A. (2017). "Context rich digital games for better learnability in the IT project management context." *Special Session on Serious Games on Computer Science Learning*. SCITEPRESS.

Nascimento, J., Da Silva, P. R. , and Samartinho, J. (2013) "Project Management Simulation Portal: Proposal, Features and Construction Process", *International Journal of Web Portals (IJWP)*, Vol. 5, No. 4, pp. 20-32.

Noe, R., Hollenbeck, J., Gerhart, B., and Wright, P. (2006), *Human Resources Management: Gaining a Competitive Advantage, Tenth Global Edition.* McGraw-Hill Education, New York.

Palotie, P., Nurkka, P., and Juvonen, S. (2017) "The quality matters: the experiences of a project manager training", *ICERI2017 Proceedings*.

Piraveenan, M. (2019) "Applications of game theory in project management: A structured review and analysis", Mathematics, Vol. 7, No. 9, pp. 858.

Pons, S., Khalilzadeh, J., Weber, M. R., and Smith, R. A. (2024) "Cultivating sustainability savvy: the role of soft skills in shaping sustainable practices", *International Hospitality Review*.

Potter, C. (1968) "The names of games managers play", *Human Resource Management*, Vol. 7, No. 4, pp. 33.

Project Management Institute (2017) *A Guide to the Body of Knowledge for Project Management (PMBOK® Guide)*, Project Management Institute, Newtown Square.

Rumeser, D., and Emsley, M. (2019) "Can serious games improve project management decision making under complexity?." *Project Management Journal* , Vol. 50, No. 1, pp 23-39.

Seager, W., Ruskov, M., Sasse, M. A. and Oliveira, M. (2011) “Eliciting and modelling expertise for serious games in project management”, *Entertainment Computing* . Vol. 2, No. 2, pp. 75-80.

Siddique, S., Ahsan, A., Azizi, N., and Haass, O. (2022) “Students’ workplace readiness: Assessment and skill-building for graduate employability”, *Sustainability*, Vol. 14, No. 3, pp 1749.

Sluka, I. (2019) “Project manager's competences game: How does it work?.” *5th IPMA SENET Project Management Conference (SENET 2019).* Atlantis Press.

Stevenson, D.H., and Starkweather, J.A. (2010) “PM Critical Competency Index: IT executives prefer soft skills”, *International Project Management Journal*, Vol. 28, No. 7, pp 663-671.

Thamhain, H. J. (1991) *Development of project management skills*, Project Management Institute.

Valderrama-Zurián, J., Aguilar-Moya, R., Melero-Fuentes, D., and Aleixandre-Benavent, R. (2015) “A systematic analysis of duplicate records in Scopus” *Journal of Informetrics*, Vol. 9, No. 3, pp 570-576.

Wyskwarski, M. (2021) “Requirements for project managers: What do job ads say?”, *Sustainability*, Vol. 13, No. 23, pp 12999.

Xue, J., Rasool, Z., Gillani, A., and Khan, A. I. (2020) "The impact of project manager soft competences on project sustainability", *Sustainability*, Vol. 12, No. 16, pp 6537.

Zoroja, J., Bach, M. P., and Miloloža, I. (2019). “Developing Leadership in a Simulation-Games Environment.” *2019 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics. IEEE.*

Zwikael, O., and Gonen, A. (2007) “Project execution game (PEG): training towards managing unexpected events”, *Journal of European Industrial Training*, Vol. 31, No. 6, pp 495-512.

Knowledge as Experience: Embodied Artistic Practices, AI, and the Governance of Intelligent Knowledge for Sustainable Innovation

Daria Podmetina[1], Anne Berthinier Poncet[2], Peter Robbins[3], Jelena Hartsenko[1], Jani Kiljala[1]

[1] Tallinn University of Technology, Tallinn, Estonia
[2] CNAM University, Paris, France
[3] Dublin City University, Dublin, Ireland

Abstract

Knowledge management (KM) research has long recognised that organisational knowledge extends beyond codified information to include tacit, experiential, and relational forms of knowing. However, contemporary organisations increasingly rely on artificial intelligence (AI) and data-driven systems that privilege abstraction, codification, and optimisation. While these systems enhance efficiency and scalability, they risk marginalising embodied and contextual knowledge that is essential for creativity, innovation, and sustainability. This paper addresses this tension by examining how organisations can develop intelligent knowledge systems while preserving the embodied foundations of human knowing. Drawing on qualitative data from 21 semi-structured interviews with innovation leaders, managers, and artists across ten European organisations, the study investigates how arts-based initiatives (ABIs) generate and mobilise embodied knowledge in organisational contexts. ABIs, defined as structured engagements embedding artistic practices within organisational processes, provide a unique empirical setting to explore knowledge that is experiential, sensory, and difficult to codify. The findings reveal that embodied knowledge constitutes a distinct form of organisational knowledge that operates through experience rather than representation. Its integration into organisational processes depends on relational translation mechanisms, including intermediaries, boundary artefacts, and collaborative practices that bridge artistic and technical domains. At the same time, AI-enabled systems emerge as a structural pressure toward abstraction, privileging codified knowledge and reinforcing optimisation logics. This creates a tension between efficiency-driven knowledge systems and the forms of knowing required for innovation under uncertainty. The paper introduces the concept of embodied knowledge governance, which explains how organisations can recognise, integrate, and sustain experiential knowledge alongside codified and AI-mediated knowledge. This framework highlights the importance of translation, interaction, and organisational conditions that preserve contextual meaning. The findings further show that embodied knowledge plays a critical role in sustainable innovation by enabling long-term, systemic thinking and expanding value perspectives beyond short-term performance. This study contributes to KM, AI, and innovation research by advancing understanding of knowledge plurality in organisations and by highlighting governance as a key mechanism for balancing intelligent systems with human creativity and sustainability.

Keywords: embodied knowledge; arts-based initiatives; knowledge governance; artificial intelligence; sustainable innovation

Paper type: Academic Research Paper

1 Introduction

The increasing integration of artificial intelligence (AI) into organizational processes is fundamentally reshaping how knowledge is created, managed, and applied (Raisch & Krakowski, 2021). Contemporary knowledge management (KM) systems are increasingly oriented toward abstraction, codification, and data-driven decision-making, privileging forms of knowledge that can be captured, stored, and recombined at scale (Nonaka and Takeuchi, 1995; Nonaka and von Krogh, 2009). While these developments enhance efficiency and scalability, they also raise critical concerns regarding the marginalisation of experiential, contextual, and embodied forms of knowing that underpin creativity, interpretation, and long-term innovation. This tension is particularly visible in innovation contexts. Organizations face growing pressure to develop intelligent knowledge systems capable of supporting decision-making under uncertainty, while simultaneously responding to sustainability challenges that require long-term, systemic thinking rather than short-term optimisation. However, dominant approaches to intelligent knowledge remain largely detached from the embodied and relational dimensions of knowledge through which meaning is constructed and shared. As one innovation leader observed, *"You know, how do you convince people of large scale risks? Not through presentations. Through the manifestation of the experience"* (B1). This highlights a fundamental limitation of abstraction-driven knowledge systems: they struggle to capture and communicate forms of knowledge that are experiential, affective, and situated. Despite extensive research on knowledge creation and transfer, existing KM literature has largely

treated embodied knowledge as a residual category of tacit knowledge, rather than as a primary knowledge asset in its own right (Nonaka and von Krogh, 2009). At the same time, emerging research on AI and intelligent knowledge systems has focused predominantly on issues of data, algorithms, and decision support, with limited attention to how these systems interact with human experiential knowing. In parallel, research on arts-based initiatives (ABIs) has demonstrated their potential to stimulate creativity, disrupt dominant cognitive frames, and support innovation capabilities (Schiuma, 2011; Berthinier-Poncet et al., 2025). However, this literature has primarily examined creativity and capability development, leaving open the question of how artistic practices shape knowledge governance in AI-enabled organizational contexts. This creates a significant research gap at the intersection of KM, AI, and innovation studies. While organizations increasingly rely on intelligent systems to manage knowledge, we still lack understanding of how embodied, experiential forms of knowledge are generated, preserved, and governed within such systems. This gap is particularly important in the context of sustainable innovation, where addressing long-term societal and environmental challenges requires integrating diverse forms of knowledge, including those that cannot be easily codified or quantified. To address this gap, this paper examines how arts-based initiatives (ABIs) enable the creation and mobilisation of embodied knowledge, and how such knowledge interacts with emerging AI-driven knowledge systems. ABIs, defined as structured and sustained engagements that embed artistic practices within organizational processes (Schiuma, 2011), provide a unique empirical setting in which experiential, affective, and relational forms of knowledge are foregrounded. By analysing how organizations engage with artistic practices, we are able to explore how embodied knowledge is produced, translated, and governed in practice.

Drawing on qualitative data from 21 interviews with innovation leaders, managers, and artists across multiple industries, the study investigates how organizations navigate the tension between embodied and abstract knowledge. The findings show that artistic practices generate distinct forms of knowledge that are not easily reducible to codified representations, but instead operate through experience, interpretation, and interaction. For example, innovation teams emphasised the importance of engaging directly with experiential artefacts rather than relying on abstract representations, as these artefacts enable shared understanding across disciplines and functions.

Building on these insights, the paper develops a conceptual framework of embodied knowledge governance, which explains how organizations can integrate experiential and AI-mediated forms of knowledge without reducing one to the other. In doing so, the paper contributes to three streams of literature. First, it extends KM research by positioning embodied knowledge as a central organizational asset. Second, it contributes to emerging debates on AI and knowledge governance by conceptualising AI as a mediating force that reshapes, rather than replaces, human knowing. Third, it advances research on sustainable innovation by showing how the governance of knowledge influences organizations' ability to address long-term societal challenges.

2 From Knowledge Management to AI Knowledge Governance and Embodied Knowledge

KM research has traditionally distinguished between explicit and tacit knowledge, emphasising processes through which knowledge can be created, shared, and applied within organisations (Nonaka and Takeuchi, 1995). Subsequent developments in KM theory have extended this perspective by recognising the importance of knowledge conversion, social interaction, and context in shaping knowledge processes (Nonaka and von Krogh, 2009). Despite these advances, dominant KM approaches remain strongly oriented toward codification, standardisation, and systematisation. This orientation reflects the increasing reliance of organisations on formal knowledge systems designed to store, retrieve, and recombine information efficiently. While such systems enable scalability and coordination, they also privilege forms of knowledge that can be abstracted from their context. As a result, experiential, affective, and embodied forms of knowing are often marginalised, even though they play a critical role in creativity, interpretation, and innovation. This tension between abstraction and experience has become more pronounced as organisations adopt digital and data-driven knowledge infrastructures.

The growing integration of artificial intelligence (AI) into organisational processes further reinforces the abstraction of knowledge. AI-enabled systems rely on data aggregation, pattern recognition, and algorithmic recombination, which favour knowledge that can be formalised and encoded (Zuboff, 2019). In this context, knowledge is increasingly treated as a resource to be optimised, raising questions about how meaning, context, and ethical considerations are preserved. Recent research on AI and knowledge governance highlights concerns regarding over-reliance on data-driven decision-making, loss of contextual understanding, and the displacement of human judgment (Verganti et al., 2020). While AI can enhance decision-making capabilities, it also introduces new challenges related to transparency, accountability, and the interpretation of knowledge outputs. In particular, AI systems tend to reproduce and recombine existing knowledge rather than generate fundamentally

new perspectives, which may limit their contribution to radical innovation. However, existing studies have largely focused on technical and managerial aspects of AI implementation, with limited attention to how AI interacts with embodied and experiential forms of knowledge (Jarrahi, 2018). As organisations increasingly rely on intelligent systems, understanding how these systems reshape knowledge processes becomes critical, particularly in contexts requiring creativity, interpretation, and long-term thinking.

In parallel, research on arts-based initiatives (ABIs) has demonstrated the capacity of artistic practices to introduce alternative modes of knowing into organisational contexts. ABIs are defined as structured and sustained engagements that embed artistic processes, practices, or actors within organisations to address managerial and innovation challenges (Schiuma, 2011). Unlike traditional problem-solving approaches, artistic practices operate through ambiguity, experimentation, and sensory engagement, enabling individuals and teams to explore new perspectives and meanings. Empirical studies show that ABIs can disrupt dominant cognitive frames, enhance creativity, and support organisational learning (Berthoin Antal, 2019; Meisiek and Barry, 2018). More recent work has emphasised their role in developing innovation capabilities and facilitating transformation within organisations (Berthinier-Poncet et al., 2025). However, this literature has primarily focused on creativity enhancement and capability development, rather than on the nature of the knowledge produced through artistic practices. Artistic practices generate forms of knowledge that are inherently embodied, relational, and experiential. These forms of knowledge are often communicated through artefacts, performances, or shared experiences rather than through formal representations. As such, they challenge conventional KM approaches and raise questions about how such knowledge can be recognised, integrated, and governed within organisations.

Sustainable innovation introduces an additional layer of complexity to organisational knowledge processes. Unlike conventional innovation, which is often driven by short-term performance and efficiency, sustainable innovation requires long-term orientation, systemic thinking, and the integration of diverse value perspectives (Schaltegger et al., 2016). Addressing sustainability challenges involves dealing with uncertainty, ambiguity, and competing priorities, making it difficult to rely solely on codified or data-driven knowledge. Recent research emphasises that sustainable innovation depends on the ability of organisations to integrate multiple forms of knowledge, including scientific, experiential, and societal knowledge (Boons and Lüdeke-Freund, 2013). This requires moving beyond purely technical or economic perspectives toward more holistic approaches that consider social and environmental implications. In this context, embodied and experiential knowledge becomes particularly important, as it enables actors to interpret complex situations, engage with stakeholders, and explore alternative futures. Despite this recognition, limited attention has been paid to how organisations govern different forms of knowledge in sustainability-oriented innovation processes, particularly in the presence of AI-driven systems.

Bringing these streams together reveals a significant gap in the literature. While KM research has acknowledged the importance of tacit knowledge, it has not fully addressed the role of embodied and experiential knowledge as organisational assets. Research on AI has focused on data-driven knowledge systems but has largely overlooked how such systems interact with non-codifiable forms of knowing. At the same time, studies on ABIs have demonstrated their creative and transformative potential but have not examined their implications for knowledge governance in AI-enabled environments. This paper addresses this gap by examining how embodied knowledge generated through artistic practices is produced, translated, and governed within organisations increasingly shaped by AI-driven knowledge systems. By focusing on the interaction between embodied, digital, and AI-mediated knowledge, the study develops a framework of embodied knowledge governance that explains how organisations can balance efficiency with meaning, and optimisation with sustainability. Figure 1 illustrates the embodied knowledge governance framework.

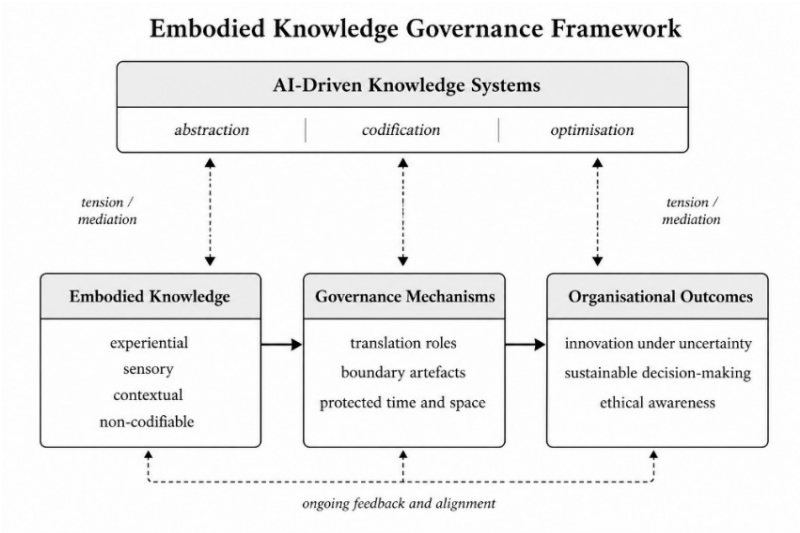


Figure 1: Embodied Knowledge Governance Framework

3 Methodology

This study adopts a qualitative, interpretive research design aimed at exploring how organisations generate, translate, and govern embodied forms of knowledge in the context of arts-based initiatives (ABIs). Given the exploratory nature of the research question and the focus on under-theorised knowledge processes, a qualitative approach is particularly appropriate for capturing rich, context-dependent insights (Gioia et al., 2013). The study builds on an empirical dataset originally developed to examine the role of ABIs in fostering innovation capabilities. While the initial research focused on dynamic capabilities, the richness of the data allows for a secondary, theoretically informed analysis centred on knowledge processes, AI-related tensions, and sustainability. This approach aligns with established practices in qualitative research, where existing datasets are re-analysed to address new research questions (Heaton, 2004). Rather than examining static organisational attributes, the study adopts a process-oriented perspective, focusing on how knowledge is produced, shared, and transformed through interaction, experience, and interpretation. This perspective is particularly relevant in contexts where knowledge cannot be fully codified and where meaning emerges through practice.

The empirical material consists of 21 semi-structured interviews conducted between November 2023 and March 2024 with both organisational actors and artists involved in ABIs (Appendix A). The sample includes: 16 interviews with organisational participants (e.g. innovation managers, R&D leaders, designers, producers, marketing directors), 5 interviews with professional artists collaborating with organisations on innovation initiatives. The organisations represented operate across a range of industries, including consumer electronics, automotive, gaming, telecommunications, consulting, luxury goods, and manufacturing. They vary in size and structure, ranging from SMEs to large multinational corporations, and are located across several European countries. This diversity of contexts enhances the analytical robustness of the study by capturing variation in how ABIs are implemented and experienced. Interviewees were selected using a purposive sampling strategy, focusing on individuals with direct experience of arts-based innovation practices. Snowball sampling was also employed to identify additional participants with relevant expertise. The inclusion of both organisational actors and artists allows for a multi-perspective understanding of how embodied knowledge is generated and interpreted. Interviews lasted between approximately 25 and 80 minutes, were conducted online, and were recorded and transcribed verbatim. The full dataset comprises over 17 hours of recorded material, resulting in more than 200 pages of transcripts. To ensure validity, transcripts were reviewed using a member-checking process, allowing participants to verify and clarify their statements.

Data analysis followed an inductive–deductive approach, combining grounded insights from the empirical material with theoretical concepts from KM, AI governance, and innovation literature (Gioia et al., 2013). The analysis proceeded in three stages: First-order coding (inductive stage): Initial open coding identified key themes emerging from the data, focusing on how interviewees described: experiential and embodied practices, knowledge sharing and communication, interactions between artistic and organisational actors, tensions between exploration and structured processes. At this stage, particular attention was given to how respondents articulated differences between experiential knowledge and formal representations. Second-order coding (theory-informed stage): Codes were then grouped into higher-order themes reflecting recurring patterns across cases. These themes included: embodied knowledge creation, translation and mediation mechanisms,

abstraction and systematisation pressures, sustainability and long-term thinking. Insights from KM and AI literature were used to interpret these patterns, particularly in relation to knowledge codification and governance. Aggregate dimensions (conceptual stage): In the final stage, these themes were synthesised into broader conceptual categories, leading to the development of the embodied knowledge governance framework. This stage involved iterative movement between data and theory to ensure conceptual coherence and theoretical contribution. Throughout the analysis, the research team engaged in iterative comparison and discussion to ensure consistency and reliability of coding decisions, following established qualitative research practices (Gioia et al., 2013).

4 Findings

The analysis reveals that arts-based initiatives (ABIs) generate and mobilise embodied forms of knowledge that interact in complex ways with abstraction-driven organisational systems. Across cases, four interrelated findings emerge concerning the nature of embodied knowledge, its translation within organisations, and its implications for AI-enabled knowledge governance and sustainable innovation.

4.1 Embodied knowledge as experiential and non-representational

The first finding shows that artistic practices generate knowledge that is experiential, situated, and difficult to codify. Rather than being expressed through formal representations such as reports or models, this knowledge is communicated through direct engagement, materialisation, and shared experience. Interviewees consistently emphasised that certain types of understanding cannot be conveyed through abstract representations alone. As one innovation leader explained: "*you know, how do you convince people of large scale risks? Not through presentations. Through the manifestation of the experience*" (B1). This highlights that embodied knowledge operates through experience rather than description, requiring individuals to engage directly with artefacts, prototypes, or situations. Such knowledge is not easily reducible to the forms required by conventional KM systems, which rely on codification and transferability. Similarly, artistic collaborators described their work as rooted in experimentation, intuition, and material engagement: "*my work takes the form of performance projects, where I generally call on materials or skills and hybridise them to come up with new perspectives and proposals…* "(A1). Together, these insights suggest that embodied knowledge is not simply tacit knowledge that has yet to be codified, but rather a distinct mode of knowing that resists abstraction. This challenges dominant KM assumptions that knowledge can always be translated into formal representations.

4.2 Translation as a relational and mediated process

The second finding shows that embodied knowledge does not automatically integrate into organisational processes; it requires active translation across domains. This translation is both relational and organisational. It involves bridging differences in language, logic, and expectations between artistic, technical, and managerial actors. In practice, this role is often performed by intermediaries such as producers or innovation leaders.

As one interviewee noted: "*it's very much a job for producers or production associates… to do this translation work… it's the ability to speak all the languages that means that, at the end of the day, your team will feel understood, whoever they are.* " (B8). Translation is therefore not simply about communication, but about reframing knowledge in ways that make it actionable across domains. At the same time, translation is not unilateral. It depends on the capacity of artists and technical actors to engage with each other's perspectives: *"you absolutely have to let the tech guys… you have to give them the space to understand that vision and amend it and make it evolve… otherwise you end up with… 'no, I can't do it'."* (B8). This highlights that translation is a co-constructed process, where knowledge evolves through interaction rather than being transferred intact. Failure to establish such translation mechanisms often results in the marginalisation of artistic insights.

4.3 AI and abstraction as pressures on embodied knowledge

The third finding concerns the role of AI and system-based knowledge architectures. While AI is not yet fully embedded in these practices, it emerges in the data as a structural pressure toward abstraction, optimisation, and standardisation. Interviewees repeatedly contrasted experiential knowledge with system-based representations, pointing to the limitations of abstraction. This tension becomes more pronounced in contexts where organisations rely on formal systems, roadmaps, and predictive tools. One respondent described how innovation processes are constrained by structured planning systems: "*Most of what designers are here to do is, let's be frank, serve the bottom line and the roadmaps of the next couple of years. And I'm afraid that is incredibly stifling to anything but incremental innovation."* (B1). Similarly, organisations were described as operating within

highly optimised systems that prioritise efficiency over exploration: *"we have optimised systems for 3 to 5 years to do things a certain way… and that actually hurts us at the end of life."* (B1).

These insights align with concerns in AI governance literature regarding over-reliance on data-driven optimisation. The findings suggest that AI-enabled systems intensify existing tendencies toward abstraction, making it more difficult for organisations to recognise and act upon embodied forms of knowledge. At the same time, interviewees implicitly distinguished artistic knowledge from system-based intelligence by emphasising its originality: *"they developed new product concepts… they have been very innovative and driven by art."* (B13). This suggests that embodied knowledge plays a complementary role to AI: rather than replacing human creativity, AI systems amplify the importance of non-recombinative, experiential insight. This interpretation is further supported by broader organisational practices. For instance, companies such as B1,2 and B10,11,12 explicitly frame AI as a tool for efficiency, optimisation, and decision-making at scale, reinforcing abstraction-driven knowledge processes. At the same time, these organisations emphasise responsible AI and sustainability goals, highlighting the need to balance data-driven systems with human-centred and contextual understanding. This pattern is consistent with industry practices, where organisations increasingly deploy AI to optimise processes and support decision-making, while recognising the need to preserve human judgment and contextual understanding.

4.4 Embodied knowledge and the reorientation toward sustainable innovation

The fourth finding connects embodied knowledge to sustainability. Across cases, artistic practices were associated with a shift from short-term optimisation toward long-term, systemic thinking. Interviewees highlighted how current innovation systems prioritise efficiency and output, often at the expense of sustainability: *"we have optimised systems …to do things a certain way…" "there is no way there is absolutely no commitment to solving this until we change the parameters…"* (B1). In contrast, artistic practices were described as enabling reflection on broader implications, including environmental and societal impact. This is particularly evident in discussions of constraints and responsibility: *"once you have to do things differently… creatives respond extremely well to those kind of constraints."* (B1). This suggests that embodied knowledge supports sustainable innovation by enabling holistic understanding of systems, fostering ethical reflection, and encouraging alternative value perspectives beyond efficiency. In this sense, ABIs do not only enhance creativity but also contribute to a reorientation of innovation processes toward sustainability and responsibility.

Taken together, the findings show that embodied knowledge is a distinct and non-codifiable form of organisational knowledge; Its integration depends on relational translation mechanisms; AI and system-based approaches create pressures toward abstraction. Embodied knowledge plays a critical role in enabling sustainable, long-term innovation. Table 2 provides illustrative quotes supporting each key finding.

5 Discussion

5.1 Reframing knowledge: from abstraction to embodied knowledge assets

This study extends knowledge management (KM) theory by reframing embodied knowledge as a first-order organisational asset, rather than a residual category of tacit knowledge. While prior research acknowledges the importance of tacit knowledge (Nonaka and Takeuchi, 1995; Nonaka and von Krogh, 2009), it has largely assumed that such knowledge can ultimately be articulated, codified, or transferred. Our findings challenge this assumption. The empirical evidence shows that artistic practices generate forms of knowledge that are inherently experiential, relational, and non-representational, and therefore resist full abstraction. Rather than being "uncaptured" knowledge, embodied knowledge operates through direct engagement, materialisation, and shared experience. This suggests a need to move beyond the traditional explicit–tacit dichotomy toward a more nuanced understanding of knowledge plurality. In this perspective, embodied knowledge is not simply a precursor to codification but constitutes a distinct epistemic mode, with its own logic of creation, validation, and transmission. Recognising this distinction is critical for organisations seeking to harness creativity and innovation in increasingly complex environments.

5.2 AI and the paradox of intelligent knowledge systems

The findings also contribute to emerging debates on artificial intelligence (AI) and knowledge governance by highlighting a fundamental paradox (Raisch & Krakowski (2021): the more organisations invest in intelligent, data-driven systems, the more they risk marginalising the forms of knowledge required for radical and sustainable innovation. AI-enabled systems rely on abstraction, standardisation, and recombination, privileging

knowledge that can be formalised and processed algorithmically (Zuboff, 2019; Verganti et al., 2020). However, the study shows that innovation—particularly under conditions of uncertainty—depends heavily on embodied, contextual, and experiential knowledge that cannot be easily encoded. Rather than positioning AI as a substitute for human cognition (Jarrahi, 2018), the findings suggest that AI functions as a mediating force that reshapes knowledge hierarchies within organisations. Specifically, AI intensifies the privileging of codified knowledge over experiential knowing, the dominance of optimisation logics over exploratory thinking, and the tension between efficiency and meaning. This leads to a reconfiguration of knowledge governance challenges. The key issue is no longer simply how to manage knowledge, but how to balance different forms of knowing within AI-enabled systems. This paradox is also reflected in organisational practice, where firms simultaneously invest in AI-driven optimisation and promote responsible AI and sustainability agendas.

5.3 Embodied knowledge governance

Building on these insights, the paper introduces the concept of embodied knowledge governance. Embodied knowledge governance refers to the set of organisational practices, structures, and mechanisms through which firms recognise, legitimise, and integrate embodied forms of knowledge, mediate their interaction with codified and AI-driven knowledge systems, and preserve their contextual and experiential integrity over time. Unlike traditional KM approaches, which focus on codification and storage, embodied knowledge governance emphasises translation rather than transfer, experience rather than representation, and interaction rather than extraction.

Table 1: Contrasting Codified and Embodied Knowledge Modalities

Dimension	Codified Knowledge	Embodied Knowledge
Form	Explicit, documented	Experiential, sensory
Transfer	Replicable	Context-dependent
AI compatibility	High	Low
Role in innovation	Incremental	Radical / exploratory
Role in sustainability	Efficiency	Meaning / responsibility

The findings suggest that effective embodied knowledge governance relies on three key mechanisms:

1. Translation mechanisms - Embodied knowledge must be actively translated across domains through intermediaries (e.g. producers, innovation leaders) who bridge artistic, technical, and managerial logics.
2. Boundary artefacts - Experiential artefacts (e.g. prototypes, demonstrations, performances) serve as mediating devices that allow embodied knowledge to be engaged without being reduced to abstract representations.
3. Protected spaces and temporal horizons - Embodied knowledge requires organisational conditions that allow exploration, including longer time horizons and protection from short-term performance pressures.

Together, these mechanisms enable organisations to integrate diverse knowledge forms without collapsing them into a single dominant logic.

5.4 Implications for sustainable innovation

The concept of embodied knowledge governance also provides new insights into sustainable innovation. Sustainable innovation requires organisations to move beyond short-term efficiency toward long-term, systemic value creation (Schaltegger et al., 2016). This, in turn, depends on the ability to integrate multiple forms of knowledge, including those that capture ethical, social, and environmental dimensions. The findings show that embodied knowledge plays a critical role in enabling such integration. By engaging individuals in experiential and relational processes, artistic practices foster reflection on long-term consequences, expand value frameworks beyond economic performance, and support the exploration of alternative futures. At the same time, the dominance of optimisation-driven systems—often reinforced by AI—can undermine these processes by prioritising efficiency and predictability. This creates a tension between intelligent knowledge systems and sustainable innovation objectives. Embodied knowledge governance provides a way to address this tension by ensuring that experiential and ethical considerations are not displaced by data-driven decision-making. In this sense, sustainable innovation is not only a matter of technological development, but also of how knowledge is governed within organisations. In this sense, embodied knowledge does not only counterbalance optimisation-driven systems but also enables proactive engagement with sustainability challenges by supporting reflexivity, ethical deliberation, and long-term orientation.

6 Conclusions

This paper examined how organisations can develop intelligent knowledge systems while preserving the embodied foundations of human creativity and judgment. Addressing a critical gap at the intersection of knowledge management, artificial intelligence, and sustainable innovation, the study explored how arts-based initiatives (ABIs) generate, translate, and govern embodied forms of knowledge within organisational contexts. Drawing on qualitative evidence from 21 interviews with managers and artists, the findings demonstrate that embodied knowledge constitutes a distinct and essential mode of knowing that cannot be fully captured by abstraction-driven systems. While organisations increasingly rely on AI-enabled knowledge infrastructures, these systems tend to privilege codified, optimised, and recombinable knowledge, thereby marginalising experiential and contextual understanding. The study shows that this creates a fundamental tension between intelligent knowledge systems and the forms of knowing required for innovation under uncertainty and sustainability challenges. To address this tension, the paper introduced the concept of embodied knowledge governance, which explains how organisations can recognise, integrate, and sustain experiential knowledge alongside codified and AI-mediated knowledge. This framework highlights the importance of translation mechanisms, boundary artefacts, and organisational conditions that enable embodied knowledge to remain actionable without being reduced to abstraction.

The study contributes to three streams of literature. First, it extends KM research by positioning embodied knowledge as a primary organisational asset, rather than a residual form of tacit knowledge. Second, it advances emerging debates on AI and knowledge governance by conceptualising AI as a mediating force that reshapes knowledge hierarchies, rather than replacing human knowing. Third, it contributes to sustainable innovation research by demonstrating that knowledge governance is a critical enabler of long-term, responsible innovation, particularly in contexts characterised by uncertainty, complexity, and competing value systems.

For practitioners, the findings suggest that organisations should move beyond a sole focus on codification and optimisation when designing knowledge systems. Instead, they should recognise the strategic value of experiential and embodied knowledge, invest in roles and structures that support cross-domain translation, create organisational conditions that allow exploration and interpretation over time, and ensure that AI-enabled systems are used in ways that augment rather than replace human judgment. In particular, organisations aiming to pursue sustainable innovation must ensure that their knowledge systems are capable of integrating ethical, contextual, and experiential considerations, rather than relying exclusively on data-driven optimisation.

This study is subject to several limitations. While it provides rich qualitative insights, it does not examine the operational deployment of AI systems directly; rather, AI emerges as an interpretive and governance concern within the data. Future research could extend these findings through longitudinal or ethnographic studies of organisations actively integrating AI into knowledge management and innovation processes. Such research would be particularly valuable in examining how embodied knowledge is transformed, preserved, or displaced over time.

7 Ethics Declaration

This study is based on qualitative interviews conducted with organisational participants and professional artists. All participants provided informed consent. Data were anonymised to ensure confidentiality.

8 AI Declaration

AI tools were used to support the structuring and editing this manuscript. All content generated was critically reviewed and revised by the authors. The final responsibility for the content and its accuracy remains with the authors.

References

Adler, N.J. (2006) 'The arts and leadership: Now that we can do anything, what will we do?', Academy of Management Learning & Education, 5(4), pp. 486–499. https://doi.org/10.5465/amle.2006.23473209

Berthoin Antal, A., Debucquet, G. and Frémeaux, S., 2019. When top management leadership matters: Insights from artistic interventions. Journal of Management Inquiry, 28(4), pp.441-457.

Berthinier-Poncet, A., Podmetina, D. and Robbins, P. (2025) 'The art of innovation: How arts-based initiatives can nurture innovation dynamic capabilities', Technovation, 148, 103337. https://doi.org/10.1016/j.technovation.2025.103337

Boons, F. and Lüdeke-Freund, F., 2013. Business models for sustainable innovation: state-of-the-art and steps towards a research agenda. Journal of Cleaner production, 45, pp.9-19.

Carlile, P.R. (2004) 'Transferring, translating, and transforming knowledge across boundaries', Organization Science, 15(5), pp. 555–568. https://doi.org/10.1287/orsc.1040.0094

Jarrahi, M.H. (2018) 'Artificial intelligence and the future of work', Business Horizons, 61(4), pp. 577–586. https://doi.org/10.1016/j.bushor.2018.03.007
Heaton, J. (2004) Reworking Qualitative Data, Sage, London.
Leonardi, P.M. and Leavell, V., 2026. Knowing enough to be dangerous: The problem of "artificial certainty" for expert authority when using AI for decision making and planning. Organization Science, 37(2), pp.516-543.
Meisiek, S. and Barry, D., 2018. Finding the sweet spot between art and business in analogically mediated inquiry. Journal of Business Research, 85, pp.476-483.
Raisch, S. and Krakowski, S. (2021) 'Artificial intelligence and management: The automation–augmentation paradox', Academy of Management Review, 46(1), pp. 192–210. https://doi.org/10.5465/amr.2018.0072
Nonaka, I. and Takeuchi, H. (1995) The Knowledge-Creating Company, Oxford University Press, Oxford.
Nonaka, I. and von Krogh, G. (2009) 'Tacit knowledge and knowledge conversion', Organization Science, 20(3), pp. 635–652. https://doi.org/10.1287/orsc.1080.0412
Tanaka, S., 2011. The notion of embodied knowledge. Theoretical psychology: Global transformations and challenges, pp.149-157.
Schaltegger, S., Hansen, E.G. and Lüdeke-Freund, F. (2016) 'Business models for sustainability', Organization & Environment, 29(3), pp. 264–289. https://doi.org/10.1177/1086026615599806
Schiuma, G., 2011. The value of arts for business. Cambridge University Press.
Schiuma, G. and Carlucci, D., 2018. Managing strategic partnerships with universities in innovation ecosystems: A research agenda. Journal of Open Innovation: Technology, Market, and Complexity, 4(3), p.25.
Verganti, R., Vendraminelli, L. and Iansiti, M. (2020) 'Innovation and design in the age of artificial intelligence', Journal of Product Innovation Management, 37(3), pp. 212–227. https://doi.org/10.1111/jpim.12523
Zuboff, S. (2019) The Age of Surveillance Capitalism, Profile Books, London.

Appendix A. List of interviewees and interviews data

Code	Role	Industry	Country, employees	Estab. year	Metadata
B1 B2	Head of design for AI & VR	Consumer electronics	Switzerland, 7.400 emp.	1981	12/02/2024 00:47:00, 9 pages
B3	Head of the HR Innovation mission	Public administration activities	France / 51.000 emp.	1859	30/01/2024 00:56:00, 11 pages
B4 B5	Marketing director & Senior brand manager	Distilliery	Ireland 19.480 emp.	1975	24/01/2024 00:25:00, 6 pages
B6	Global VP of R&D	Telecom / R&D	Ireland, 10.000 emp.	1925	23/02/2024 01:03:00, 13 pages
B7	CEO of the Irish branch	Market research	Ireland, 19.000 emp.	1975	30/01/2024 01:10:00, 17 pages
B8	Executive producer	Video Game industry	France, 21.000 emp	1986	2/02/2024 00:52:00, 11 pages
B9 A2	CEO, Photographer & innovation manager (the artist residence)	Manufacturing	France, 250 emp.	1923	14/02/2024 00:43:00 11 pages 24/01/2024 01:25:00 18 pages
B10 B11 B12	Product manager. Product innovation. Interior product manager	Automotive	France/ 31.000 emp.	1898	18/12/2023 01:15:00 12 pages,01/02/2024 00:30:00 5 pages, 05/01/2024 00:50:00 9 pages
B13 B14	Former Head of Innovation Senior platform manager	Luxury goods	Austria, about 29 000 emp.	1895	23/12/2023 01:00:00 10 pages,12/02/2024 01:00:00, 14 pages
B15	Consultant in innovation, design and sustainability	Luxury	UK		07/03/2024 00:58:00 11 pages
B16	Head of Knowledge Transfer Institute	R&D	Austria, 520 emp.	1688	06/02/2024 00:29:00 6 pages
A1	Food Artist + designer / Innovation and Futures consultant	Arts	France		26/01/2024 01:03:18 13 pages
A3	Sculptor (artist-in-residence)	Arts	France		22/02/2024 01:14:00, 16 pages
A4	Consultant in change management and Art thinking	Arts	Spain		08/02/2024 01:16:00, 18 pages
A5	Visual artist	Arts	France		25/03/2024 01:18:00, 18 pages

Smart Knowledge for Sustainable Artisanal Production in the Integration of Science, Technology and Local Knowledge in Rural Communities in Guanajuato

Fabiola Colmenero Fonseca[1], Daniel Ruiz López[2,3], Ramiro Rodríguez Pérez[3], Lauren Yolanda Gómez Zamorano[4], Amparo Borrell[1]

[1] Universitat Politècnica de València. Valencia, Spain
[2] Universidad Autónoma de Guadalajara, Zapopan, Jalisco, Mexico
[3] Tecnológico Nacional de México. Jalisco, Mexico
[4] Universidad Autónoma de Nuevo León, Monterrey, Mexico

Abstract

This study identifies and empirically validates a microfoundational mechanism based on the Intelligent Knowledge Management Model (MIGC) to explain how knowledge systems mediated by artificial intelligence (AI) enable the twin transition in artisanal production contexts. Although the literature has extensively analyzed the relationship between digitalization and sustainability, a critical gap persists in the understanding of the microfoundational processes through which these transformations emerge in environments of low technological intensity and strong cultural roots.

Through a combined approach of design science and participatory action research, an intervention in a rural ceramics system in Guanajuato, Mexico, is implemented and evaluated. The model integrates AI, energy monitoring and artisanal knowledge within an architecture aimed at the recombination of knowledge and the development of organizational capacities. The data were collected over 12 months through energy audits, production records and qualitative fieldwork.

The results show a 19.6% reduction in energy intensity and a 14.2% reduction in production defects, along with improvements in process stability, traceability and decision-making. Beyond these results, the study identifies a recursive microfoundational mechanism in which AI does not act as an automation tool, but as a cognitive infrastructure that facilitates the recombination of tacit and explicit knowledge, activating the reconfiguration of organizational capacities and generating sustainable results.

This finding allows us to reconceptualize the twin transition as a socio-cognitive process based on the articulation of knowledge, rather than as a direct consequence of technological adoption. Likewise, the study demonstrates that dynamics of innovation ecosystems can emerge at microterritorial scales and that their effectiveness depends on the interaction between technology, agency and knowledge structures.

Overall, the results establish that sustainability in contexts of low technological intensity does not depend on intensive automation, but on the design of knowledge architectures that integrate artificial intelligence, local knowledge and organizational agency, opening new lines of research on the scalability of these processes.

Keywords: Artificial intelligence; Knowledge recombination; Knowledge management; Twin transition; Microfoundations; Artisanal production; Sustainable production; Innovation ecosystems

Paper type: Academic Research Paper

1 Introduction

The climate crisis and accelerated digitalisation are shaping a double structural transformation of contemporary economies. However, the distribution of digital and energy skills is profoundly unequal. While industrial ecosystems are moving towards decarbonization supported by smart technologies, rural production systems, particularly artisanal ones, remain lagging behind in terms of digital infrastructure, energy autonomy and access to technological innovation (Aklilu et al., 2025; Zhang et al., 2024). Recent literature has conceptualized this process as a "twin transition", understood as the coordinated integration between digital transformation and environmental sustainability (Ben Youssef, 2025). However, empirical evidence has predominantly concentrated on large organizations, smart cities, and high-tech manufacturing sectors, leaving unexplored how this transition can be operationalized in peripheral economies based on tacit knowledge and community organizational structures, as well as in regional green innovation ecosystems (Zhang et al., 2024). This limitation reveals a relevant theoretical tension: although the twin transition is often framed as a process driven by technological adoption, the underlying mechanisms that enable its outcomes remain insufficiently explained, particularly at the microfoundational level (Breiter et al., 2024). Consequently, little is known about how knowledge systems allow transformation processes to be activated in traditional productive contexts with limited resources.

In parallel, innovation ecosystem theory has evolved towards approaches that emphasize knowledge flows, dynamic capacities, territorial governance, and multilevel relationships between public, private, and territorial actors (Costa & Moreira, 2022; Ferreira et al., 2023; West & Olk, 2024; Shin et al., 2023).However, these perspectives maintain a macrostructural emphasis and do not explain how, at the micro level, processes of reconfiguration of capacities are activated when technologies such as artificial intelligence are introduced in artisanal production contexts. In this context, the central problem of this study is the absence of microfoundational explanations on how knowledge systems mediated by artificial intelligence reconfigure capacities in artisanal production environments. In particular, key questions persist, such as the following: how are energy capacities reconfigured, how are tacit and scientific knowledge recombined? And how is cultural identity preserved by avoiding processes of substitutive automation? Based on this, this study addresses the following research question: How can knowledge systems mediated by artificial intelligence enable the twin transition in rural artisanal productive ecosystems? Based on the above, the present work makes three main contributions:

First, it identifies a microfoundational mechanism that explains how artificial intelligence enables the reconfiguration of capacities at the micro level in artisanal production systems.

Second, it reconceptualizes the twin transition as a process of knowledge recombination, rather than as a direct result of technological adoption.

Third, it incorporates technological agency as a key explanatory dimension, showing that the appropriation and meaningful use of digital tools constitutes a structural condition for sustainable transformation.

2. Theoretical Framework

2.1 Knowledge ecosystems and micro-foundations

Innovation ecosystems are conceptualized as interdependent configurations of actors that co-create value through the exchange and recombination of knowledge (Öberg & Alexander, 2019; Thomas & Ritala, 2025). Likewise, these ecosystems include institutional relationships, public policies and territorial dynamics that influence innovative performance (Costa & Moreira, 2022; Shin et al., 2023). More recently, the approach has incorporated the notion of microfoundations, recognizing that dynamic capacities are activated through specific cognitive and organizational processes (Breiter et al., 2024). In rural contexts, knowledge functions as a territorial common good, deeply incorporated into productive practices and social relations. Dynamic capabilities (sensing, harnessing and reconfiguring) allow structures to be adapted to environmental pressures. However, in artisanal systems these capacities are not formalized; they depend on tacit knowledge and accumulated experience.

2.2 Twin transition in traditional production systems

Digital transformation is aimed at optimizing processes through computer technologies. The twin transition involves a socio-technical reconfiguration that integrates digitalization and environmental sustainability in a structural way (Ben Youssef, 2025). In traditional systems, tensions emerge:

- risk of cultural homogenization.
- external technological dependence.
- displacement of local knowledge.

Therefore, the twin transition requires contextualized mediation rather than simple technological incorporation (Zhang et al., 2024).

2.3 AI as a cognitive mediator

AI can be understood as augmented intelligence infrastructure. Instead of automating, it amplifies human capacity for analysis and decision-making. To this end, two mechanisms are central:

- Digital twins: energy and process simulation.
- Knowledge recombination: structured integration of heterogeneous knowledge (Marino & Quatraro, 2023; Nan, 2024).

In artisanal environments, AI facilitates correlations between material intuitions and technical data, activating processes of reconfiguration of capabilities.

2.4 Women's technology agency

The theory of self-determination states that autonomy, competence and belonging sustain intrinsic motivation (Vansteenkiste et al., 2006; Abós et al., 2019). In rural communities led by women, technology can strengthen cohesion and resilience if it is integrated in a participatory manner (see Figure 1).

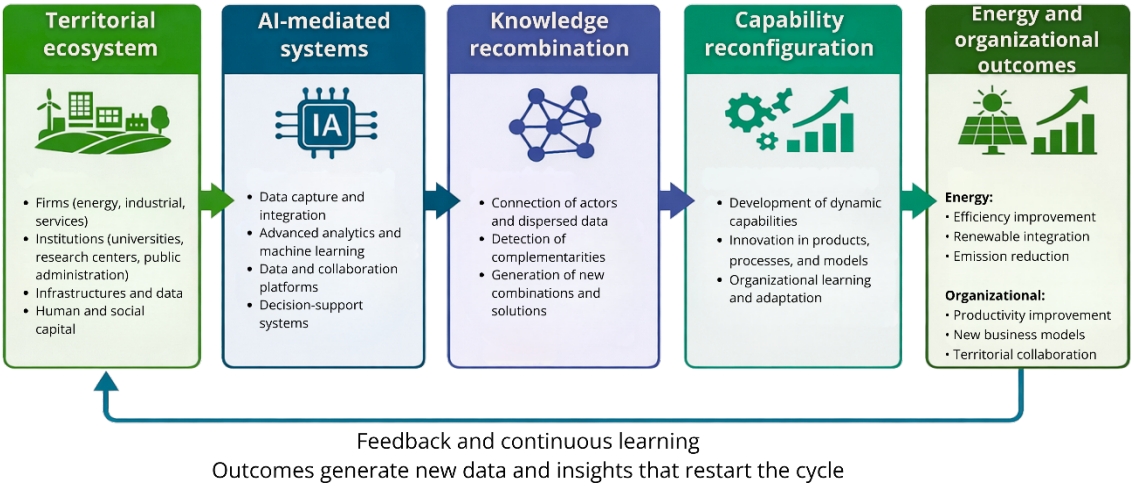


Figure 1. Microfoundational mechanism of AI-mediated twin transition

3 Methodology

3.1 Research design

The study adopts the Design Science Research (DSR) approach to the design, implementation and evaluation of an organizational artifact: the Intelligent Knowledge Management Model (MIGC). This approach is relevant when the goal is to solve a complex practical problem by building and iteratively validating a theoretically grounded artifact. The process followed the classic phases of DSR:

- identification of the problem,
- definition of artifact objectives,
- design and development,
- demonstration in real context,
- empirical evaluation,
- Communication of results.

In order to strengthen the contextual validity and organizational appropriation of the model, the design was integrated with a participatory action research approach, allowing co-creation processes with productive actors and iterative adjustment cycles based on continuous feedback. Beyond its practical orientation, DSR allows for the identification of generative mechanisms, making it suitable for microfoundational theorizing (Breiter et al., 2024).

3.2 Case selection

A women-led artisanal ceramics company in Dolores Hidalgo, Guanajuato, Mexico, was selected through purposive sampling. The choice responded to theoretical-analytical criteria:

- High energy intensity in cooking processes.
- Dependence on traditional ovens with high thermal inefficiency.
- Predominance of intergenerational tacit knowledge.
- Limited prior digitalization.
- Explicit willingness to participate in experimental processes.

The case constitutes a critical environment to analyze the twin transition (digital and energy) in traditional production systems (Ben Youssef, 2025; Zhang et al., 2024).

3.3 Data collection

The intervention lasted 12 months, combining quantitative and qualitative methods to ensure triangulation and analytical robustness. To this end, the following data were collected:

- Systematic energy audits (kWh per production batch).
- Longitudinal measurement of defect rate (%).
- Structured digital recording of process parameters.
- 18 semi-structured interviews with owners and operators.
- Participant observation during complete production cycles.
- Documentary analysis of routines and historical records.

Quantitative data were analyzed by pre-post-intervention comparison, while qualitative data were thematically coded to identify microfundamentals of capacity reconfiguration as shown in Table 1 below.

Table 1: Five-phase intervention model.

Phase	Objective	Dynamic Capability On	Indicators
Diagnosis	Establish energy and organizational baseline	Sensing	kWh/lot; defect rate
Recombination	Experimental redesign of cooking parameters	Utilization	% reduction in defects
AI Mediation	Implementation of data-driven decisional support	Reconfiguration	Digital logs; Assisted decisions
Solar integration	Incorporation of photovoltaic energy into the process	Reconfiguration	% solar energy used
Institutionalization	Formalization of routines and protocols	Stabilization	Documented protocols; Internal Adoption

4 Results

4.1 Reconfiguration of energy capacities

This section presents the empirical results derived from the intervention, without theoretical interpretation. The quantitative results show a significant reconfiguration of energy and operational capacities after the implementation of the Intelligent Knowledge Management Model. Energy intensity was reduced by 19.6%, thermal variability by 17%, and the proportion of defective batches by 14.2% compared to baseline. Additionally, a decrease in the dispersion of the cooking cycles was observed, as well as greater stability in the thermal parameters recorded during the production process. This stabilization allowed to reduce critical deviations and improved consistency between batches.

Table 2: Comparative pre- and post-intervention indicators.

Indicator	Baseline	Pilot	Variation
Energy intensity	100%	80.4%	−19.6%
Defective batches	100%	85.8%	−14.2%
Thermal variability	100%	83%	−17%

Together, the results show simultaneous improvements in energy efficiency, product quality, and process control.

4.2 Hybridization and recombination of knowledge

A systematic process of integration between artisanal knowledge and digitized technical parameters was observed. During the intervention, the participants began to use data records to complement their decisions based on experience and the implementation of digital tools made it possible to correlate variables with specific productive results such as:

- Temperature.
- Cooking time.
- Energy consumption.

This correlation facilitated the identification of previously unformalized patterns, consistent with recent approaches that explain how knowledge recombination increases innovation capacities when heterogeneous cognitive domains are integrated (Marino & Quatraro, 2023; Nan, 2024; Wang et al., 2023). Likewise, a progressive transition from decisions based exclusively on intuition to decisions based on data was recorded,

keeping practical experience as a central reference. This process gave rise to a hybrid workspace where tacit knowledge and explicit knowledge were used in a complementary way.

4.3 Augmented decision architecture

The incorporation of energy dashboards modified the dynamics of decision-making within the organization. The performance indicators were integrated into weekly productive meetings, structuring discussions around shared metrics, where it was observed:

- Increased frequency of adjustments in production parameters.
- Reduction of uncertainty associated with cooking processes.
- systematic documentation of decisions.
- Greater operational traceability.

In addition, the available information made it possible to anticipate deviations and make preventive, rather than corrective, adjustments. This change was reflected in greater coordination among the participants and in better planning of production cycles. These findings dialogue with the notion of dynamic capacities and knowledge ecosystems (Breiter et al., 2024; Ferreira et al., 2023).

4.4 Technology Agency

The qualitative analysis shows an increase in the perception of autonomy and technical competence among the participants (Abós et al., 2019). The interviews show that the direct use of digital tools allowed the operators to:

- Interpret data independently.
- Actively participate in decision-making.
- Engage in controlled experimentation.

Likewise, a strengthening of collective decision-making was observed. In addition, the weekly meetings incorporated indicators as a central element of discussion, facilitating agreements based on shared information.

Table 3: Qualitative evidence of reconfiguration and agency.

Dimension	Evidence	Interpretation
Autonomy	Independent use of dashboards	Technological appropriation
Competition	Experimenting with metrics	Active learning
Collective decision	Meetings with shared indicators	Internal coordination

Overall, the qualitative results indicate a greater appropriation of digital tools, as well as an evolution in the dynamics of interaction and decision-making within the production system.

5 Discussion

This study challenges the dominant assumption that digital transformation automatically precedes sustainability outcomes. The findings show that, in artisanal production contexts, sustainability does not emerge from technological adoption per se, but from internal processes of knowledge recombination mediated by artificial intelligence (Marino & Quatraro, 2023; Nan, 2024). In this sense, the twin transition is configured as a socio-cognitive process rather than as an exclusively technological phenomenon.

Unlike approaches focused on digital investment or infrastructure deployment, the results show that transformation is activated when AI is integrated into organizational routines as a cognitive mediator, enabling new ways of interpreting, articulating and using existing knowledge.

5.1 From digital infrastructure to cognitive mediation

The literature on twin transition has tended to conceptualize digitalization as an exogenous enabler of sustainability, mainly at macro levels (Aklilu et al., 2025; Ben Youssef, 2025). This study challenges that premise by demonstrating that the transformative impact of AI depends on its internalization into concrete organizational practices. Instead of operating as an automation tool, AI functioned as cognitive infrastructure, reconfiguring the way the organization processes information and makes decisions. This finding shifts the focus from technological adoption to cognitive integration, suggesting that digitalization only generates value when it modifies the internal logics of action. In this sense, the results broaden the focus on dynamic capabilities by showing that the transition to more advanced stages does not depend solely on the accumulation of digital

resources, but on the ability to incorporate them into recurrent adaptive processes (Breiter et al., 2024). Empirical evidence indicates that the key is not to "have data", but to use it systematically in decision-making. Likewise, the study qualifies the literature on digital ecosystems by showing that technology does not replace pre-existing knowledge structures, but rather reconfigures them (Costa & Moreira, 2022; Ferreira et al., 2023; Liu et al., 2025).

In this context, it is important to highlight that AI acted as a translation mechanism that made it possible to articulate tacit and explicit knowledge, reducing cognitive friction and enabling new combinations.

5.2 The microfoundational mechanism

The main contribution of the study lies in the identification of a microfoundational mechanism that explains how AI translates into sustainable outcomes. Unlike linear models, the findings suggest a recursive process in which technology and knowledge co-evolve. AI does not generate direct impact; Its effect is produced through the recombination of heterogeneous knowledge. This recombination activates organizational reconfiguration processes, which in turn generate improvements in energy and production performance.

Sustainability, therefore, is not an immediate result of digitalization, but an emerging consequence of changes in organizational capabilities. This finding extends the literature on knowledge recombination by demonstrating that value does not lie only in the accumulation of knowledge, but in its ability to articulate different domains and complementary cognitive structures (Nan, 2024; Wang et al., 2023). In addition, AI plays a central role in reducing the cognitive distance between tacit knowledge and formal metrics, facilitating more controlled experimentation processes (Marino & Quatraro, 2023). From the perspective of innovation ecosystems, the case shows that dynamics of co-creation, exchange and articulation of knowledge are not exclusive to macro environments, but can emerge in small and territorially situated configurations (Costa & Moreira, 2022; Shin et al., 2023; Thomas & Ritala, 2025; West & Olk, 2024). This suggests that ecosystems should not be understood only as external structures, but also as internal processes of cognitive coordination.

5.3 Motivation and agency as catalysts

A distinctive contribution of the study is the incorporation of the motivational dimension as part of the twin transition process. The results show that the transformation not only implied technical changes, but also a reconfiguration in the perception of autonomy and competence. Evidence suggests that technological appropriation strengthened participants' agency, consistent with self-determination theory postulates about autonomy, competence, and intrinsic motivation (Vansteenkiste et al., 2006; Abós et al., 2019). This finding contrasts with views that associate digitalization with labor substitution, showing instead its potential as an empowerment mechanism. In this context, knowledge recombination was accompanied by motivational recombination.

The integration of digital tools not only expanded technical capabilities, but also reinforced confidence in decision-making. This indicates that the effectiveness of the twin transition depends on the alignment between technological infrastructure and human agency, but also as internal processes of cognitive coordination.

5.4 Border conditions

The results allow us to identify conditions that delimit the scope of the proposed mechanism. First, organizational cohesion facilitated the integration of data into collective decision-making processes. The existence of trust and internal communication allowed digital information to be translated into coordinated action.

Second, participatory governance was key to the appropriation of the tools. The incorporation of indicators in deliberative spaces made it possible to transform data into concrete decisions.

Third, technological appropriation was a determining factor.

Evidence suggests that, without internalization by users, AI would have remained an external resource with no significant impact.

These conditions indicate that the mechanism identified is not automatic or universal, but contingent on specific organizational configurations.

5.5 Theoretical implications

The results allow us to identify conditions that delimit the scope of the proposed mechanism. First, organizational cohesion facilitated the integration of data into collective decision-making processes. The

existence of trust and internal communication allowed digital information to be translated into coordinated action.

Second, participatory governance was key to the appropriation of the tools.

The incorporation of indicators in deliberative spaces made it possible to transform data into concrete decisions.

Third, technological appropriation was a determining factor.

Evidence suggests that, without internalization by users, AI would have remained an external resource with no significant impact.

These conditions indicate that the mechanism identified is neither automatic nor universal, but contingent on specific organizational configurations (see Figure 2).

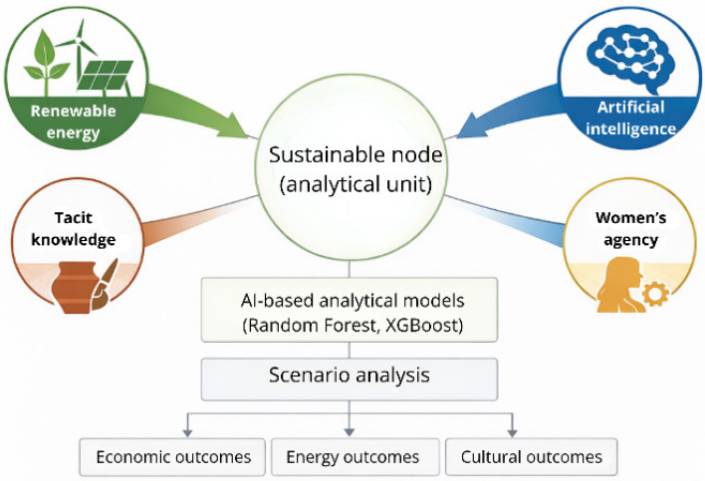


Figure 2. Smart node for micro-territorial sustainability.

Taken together, these findings suggest that the twin transition should be theorized not as a technology-driven transformation, but as a microfoundational knowledge-mediated process in which artificial intelligence functions as a cognitive infrastructure that enables the recursive reconfiguration of capabilities.

6 Implications

This study allows us to derive relevant implications at the organizational, systemic and conceptual levels. At the organizational level, the implementation of AI in artisanal environments is viable when an incremental and participatory approach is adopted. The impact does not depend on technological sophistication, but on the ability to integrate AI into productive routines and decision-making processes.

At the systemic level, small-scale productive units can operate as innovation microecosystems, consistent with recent approaches to regional and territorial ecosystems and public innovation policies (Costa & Moreira, 2022; Ferreira et al., 2023; Shin et al., 2023; Zhang et al., 2024). The progressive integration of digital tools facilitates the articulation between tacit knowledge and structured data. In intervention design, effectiveness depends on the technological appropriation by the actors. Finally, the scalability of the twin transition depends on:

- AI integration.
- Recombination of knowledge.
- Organizational agency.

7 Limitations

This study has limitations that should be considered. First, it is a unique case study in the pilot phase. Second, the timing of the study is limited, making it impossible to evaluate long-term effects. Third, although agency strengthening and technological appropriation were documented, these dimensions were mainly addressed through qualitative evidence.

Further studies could incorporate psychometric instruments to measure:

- Autonomy.
- Jurisdiction.

· Motivation.

Likewise, the proposed mechanism may be sensitive to specific contextual conditions.

8 Conclusions

This study demonstrates that the twin transition is not a direct result of technological adoption, but a micro-foundational process of a socio-cognitive nature mediated by knowledge systems. Empirical evidence shows that artificial intelligence does not generate sustainable impact directly, but acts as a cognitive infrastructure that enables the recombination of knowledge and the reconfiguration of capabilities (Liu et al., 2025; Nan, 2024). The main contribution lies in the identification of a recursive mechanism through which the integration between tacit and explicit knowledge activates dynamics of organizational reconfiguration (Marino & Quatraro, 2023; Nan, 2024). Likewise, the study demonstrates that processes typical of innovation ecosystems can emerge at microterritorial scales (Thomas & Ritala, 2025; West & Olk, 2024). In this framework, the twin transition is redefined as an emerging, contingent and potentially scalable process.

9 Acknowledgments

This publication is part of the Scientific and Humanistic Research Projects in Strategic Axes 2025, Budget Program F003 (PpF003) "National Strategic Programs of Science, Technology and Linkage with the Social, Public and Private Sectors" Project PEE-2025-C-6 of the Secretariat of Science, Humanities, Technology and Innovation SECIHTI. / This publication is part of the Scientific and Humanistic Research Projects in Strategic Axes 2025, Budget Program F003 (PpF003) "National Strategic Programs of Science, Technology and Linkage with the Social, Public and Private Sectors" Project PEE-2025-C-6 of the Secretariat of Science, Humanities, Technology and Innovation SECIHTI.

10 Ethics Declaration

This study was conducted in accordance with the ethical principles established for academic research. It did not involve the participation of humans or animals. In case of data use, these were obtained from public sources or processed anonymously, guaranteeing confidentiality and compliance with current regulations.

11 AI Declaration

The authors state that artificial intelligence tools were used to support the writing, organization, and revision of the text. However, the research design, data analysis, and conclusions were developed entirely by the authors, who assume full responsibility for the content.

12 Conflict of Interest Statement

The authors declare no conflicts of interest in relation to the publication of this article.

References

Abós, Á., Julián, J. A., Abarca-Sos, A., Aibar, A., & García-González, L. (2019). Does the level of motivation of physical-education teachers matter in terms of job satisfaction and emotional exhaustion? A person-centred examination based on self-determination theory. *International Journal of Environmental Research and Public Health*, *16*(16), 2839. https://doi.org/10.3390/ijerph16162839

Aklilu, A., Dussaux, D., & Verrier, A. (2025). *Quantifying digital innovation for the twin transition: Historical trends, current landscape and links with environmental innovation* (OECD Environment Working Papers No. 261). OECD Publishing. https://doi.org/10.1787/8cc4cff0-en

Ben Youssef, A. (2025). Twin transition: Digital transformation pathways for sustainable innovation. *Sustainability*, *17*(21), 9491. https://doi.org/10.3390/su17219491

Costa, J., & Moreira, A. C. (2022). Public policies, open innovation ecosystems and innovation performance: Analysis of the impact of funding and regulations. *Journal of Open Innovation: Technology, Market, and Complexity*, *8*(4), 210. https://doi.org/10.3390/joitmc8040210

Liu, H., Liu, Y., Huang, T., & Zhang, H. (2024). The effects of entrepreneurial ecosystems, knowledge management capabilities, and knowledge spillovers on international open innovation. *R&D Management*, *53*(2), 322–338. https://doi.org/10.1111/radm.12569

Liu, H., Liu, Y., Huang, T., & Zhang, H. (2025). A study of the influence mechanism of enterprise innovation ecosystems on digital innovation capabilities. *Sustainability*, *17*(13), 5837. https://doi.org/10.3390/su17135837

Nan, D. (2024). The influence of proximity and knowledge base on recombination innovation in R&D collaboration. *PLOS ONE*, *19*(2), e0298735. https://doi.org/10.1371/journal.pone.0298735

Öberg, C., & Alexander, A. T. (2019). The openness of open innovation in ecosystems: Integrating innovation and management literature on knowledge linkages. *Journal of Innovation & Knowledge*, *4*(4), 211–218. https://doi.org/10.1016/j.jik.2017.10.005

Shin, H., Kim, K., & Lee, J. (2023). Inventors, firms and localities: Insights into the nexus that forms and alters the evolution of regional knowledge spaces. *Cambridge Journal of Regions, Economy and Society*, *16*(3), 511–528. https://doi.org/10.1093/cjres/rsad019

Teece, D. J., Peteraf, M., & Leih, S. (2024). Dynamic capabilities for the twin transformation climb: A capability maturity model. *Information Systems Frontiers*, *26*, 2205–2226. https://doi.org/10.1007/s10796-024-10520-y

Thomas, L. D. W., & Ritala, P. (2025). Open innovation ecosystems. In M. Dodgson, D. M. Gann, & N. Phillips (Eds.), *Elgar encyclopedia of innovation management*. Edward Elgar Publishing. https://doi.org/10.4337/9781035306459.00039

Vansteenkiste, M., Lens, W., & Deci, E. L. (2006). Intrinsic versus extrinsic goal contents in self-determination theory: Another look at the quality of academic motivation. *Educational Psychologist*, *41*(1), 19–31. https://doi.org/10.1207/s15326985ep4101_4

Wang, C.-J., Yan, L., & Cui, H. (2023). Unpacking the essential tension of knowledge recombination: Analyzing the impact of knowledge spanning on citation counts and disruptive innovation. *Preprints*. https://doi.org/10.20944/preprints202506.0739.v1

West, J., & Olk, P. (2024). Open innovation, innovation in ecosystems, and innovation beyond the firm. In H. Chesbrough, W. Vanhaverbeke, & J. West (Eds.), *The Oxford handbook of open innovation* (pp. 287–307). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780192899798.013.18

Zhang, K., Wen, Y., & Wu, Y. (2024a). How digital innovation ecosystems facilitate low-carbon transformation of the economy based on a dynamic qualitative comparative analysis. *Sustainability*, *16*(22), 9962. https://doi.org/10.3390/su16229962

Zhang, K., Wen, Y., & Wu, Y. (2024b). How does the digital innovation ecosystem enable green regional development? A dynamic QCA study in China. *Systems*, *12*(12), 551. https://doi.org/10.3390/systems1212551

Intelligent Knowledge Ecosystems for the Dual Transition, Curricular Innovation, and Sustainable Production in Artisan Communities of Mexico

Juan Francisco Palomino Bernal[1], Fabiola Colmenero Fonseca[2], Rosa María López Salinas[1], Elsa Ivette Aldana Ornelas[1], Javier Cárcel Carrasco[2]

[1] Tecnológico Nacional de México, Jalisco, Mexico

[2] Universitat Politècnica de València, Valencia, Spain

Abstract

The dual transition, understood as the convergence of digital transformation and environmental sustainability, has emerged as a central strategic paradigm for territorial development and innovation policy. However, existing scholarship has predominantly focused on urban, industrial, and technologically mature environments, leaving rural and artisanal production systems underexplored. This gap is theoretically significant because conventional innovation ecosystem frameworks insufficiently explain how innovation emerges in contexts where tacit, culturally embedded knowledge constitutes the dominant productive resource. To address this limitation, this study introduces the concept of Intelligent Knowledge Ecosystems (IKEs), defined as socio-technical and territorially embedded configurations in which collective intelligence, knowledge recombination, and accessible digital-environmental capabilities interact to generate adaptive, inclusive, and sustainable innovation processes. Unlike conventional innovation ecosystems, which frequently emphasize formal institutional arrangements and market-oriented value creation, IKEs position tacit community knowledge as a strategic innovation asset and conceptualize technology as an enabler of collective learning rather than as an exogenous driver of transformation. Methodologically, the research adopts a participatory action research design implemented over a 12-month intervention in an artisan ceramic production community in Dolores Hidalgo, Guanajuato, Mexico. The study involved 48 participants, including artisans, university students, researchers, and institutional actors. Data were collected through semi-structured interviews, co-creation workshops, and field observations, and analyzed using thematic coding and process tracing to identify causal mechanisms underpinning ecosystem emergence and innovation dynamics. The findings indicate that innovation in this context does not emerge through isolated technological adoption, but through structured processes of knowledge recombination between traditional artisanal expertise and contemporary digital and sustainability-oriented tools. These interactions improved productive efficiency, supported more sustainable production practices, strengthened local capabilities, and reinforced cultural identity. The study contributes in three principal ways. First, it advances innovation ecosystem theory by proposing the IKE construct as a more inclusive and context-sensitive framework. Second, it provides empirical evidence from a traditionally underrepresented setting. Third, it offers a transferable model for integrating universities, communities, and accessible technologies in support of territorially grounded dual transition strategies.

Keywords: Dual transition, innovation ecosystems, knowledge recombination, intelligent knowledge ecosystems (IKEs), collective intelligence, participatory action research

Paper type: Academic Research Paper

1 Introduction

The dual transition, referring to the convergence between digital transformation and environmental sustainability, has become a defining strategic framework in contemporary economic and territorial development policy (Kovacic et al., 2024; Quirico & Baber, 2024). This paradigm reflects the growing recognition that digital technologies and sustainability objectives should not be addressed as separate domains, but as interconnected drivers of systemic transformation. In this context, digital technologies are increasingly positioned not merely as instruments of automation, but as enablers of new production models characterized by greater adaptability, resource efficiency, and environmental responsiveness (Chen et al., 2023; Ying & Jin, 2024).

At the territorial level, the dual transition requires a reconceptualization of innovation processes beyond conventional industrial and urban ecosystems. Contemporary scholarship increasingly acknowledges the importance of collaborative structures involving firms, institutions, governments, and civil society in generating innovation capacity (Oliveira & Rua, 2025; Pattinson et al., 2023). However, these frameworks remain strongly biased toward formal, technologically advanced, and market-driven environments, thereby marginalizing contexts where innovation emerges through informal, tacit, and culturally embedded practices. This omission is particularly problematic in artisan communities, where production systems are deeply rooted in inherited

knowledge, empirical experimentation, and collective cultural memory. Such contexts represent complex socio-technical environments in which tacit knowledge is not peripheral, but foundational to productive activity and local resilience (Yang, 2024). Despite their relevance, these systems have historically remained peripheral within innovation ecosystem literature, limiting theoretical understanding of how digital and sustainability transitions may unfold in such environments (Heaton & Min, 2025). At the same time, recent scholarship has emphasized knowledge recombination as a central mechanism of innovation, highlighting how the integration of previously distinct knowledge domains can generate novel solutions and adaptive capabilities (Xiao et al., 2022; Yu & Yu, 2024). Yet, existing research has largely examined recombination in formal R&D-intensive settings, leaving important questions unanswered regarding how such mechanisms operate in low-technology, territorially embedded production systems.

This theoretical gap is compounded by conceptual ambiguity regarding the role of technology in innovation ecosystems. In dominant perspectives, technology frequently appears as an exogenous catalyst that accelerates transformation through efficiency gains or infrastructural modernization. However, this framing underestimates the extent to which innovation may emerge through socially mediated learning processes, collective experimentation, and contextual adaptation rather than through technological sophistication alone.

To address these limitations, this study proposes the concept of Intelligent Knowledge Ecosystems (IKEs) as a theoretical framework capable of explaining innovation dynamics in territorially embedded, culturally grounded, and resource-constrained environments. Rather than conceptualizing innovation as a product of formal institutional architecture alone, IKEs emphasize the interaction between collective intelligence, knowledge recombination, and accessible technological mediation. Empirically, the study examines a participatory intervention conducted in an artisan ceramic production community in Guanajuato, Mexico, where university actors, artisans, and institutional stakeholders collaboratively engaged in sustainability-oriented innovation processes. This case offers a relevant setting to analyze how dual transition mechanisms can emerge in contexts typically excluded from mainstream innovation ecosystem discourse. The research addresses the following question: How do Intelligent Knowledge Ecosystems enable dual transition processes in rural artisan production systems? This study makes three primary contributions.

First, it introduces and theoretically defines the concept of Intelligent Knowledge Ecosystems as an extension of innovation ecosystem theory.

Second, it empirically demonstrates how knowledge recombination and collective learning operate as mechanisms through which innovation emerges in culturally embedded production contexts.

Third, it provides a practical and transferable framework for linking higher education institutions, local communities, and sustainability-oriented technologies in support of inclusive territorial development.

2 Theoretical Framework

2.1 Innovation ecosystems and the limits of dominant perspectives

Innovation ecosystems are commonly understood as dynamic networks of interdependent actors that collaboratively create value through the exchange of knowledge, resources, and capabilities (Oliveira et al., 2025; Pattinson et al., 2023). These ecosystems are generally characterized by distributed innovation processes, open collaboration, institutional interdependence, and adaptive governance structures that facilitate collective value creation beyond traditional firm-centered models. Public policy has also been increasingly recognized as a significant driver of ecosystem performance, particularly where institutional arrangements and innovation governance influence knowledge exchange and collaborative competitiveness (Costa & Moreira, 2022). Within this literature, knowledge recombination has emerged as a particularly important mechanism for innovation generation, referring to the integration of existing knowledge components into novel configurations capable of producing adaptive and competitive outcomes (Xiao et al., 2022; Luo, 2022). Diverse actor configurations are frequently associated with enhanced innovation capacity because heterogeneous knowledge domains create conditions for experimentation, iterative learning, and organizational adaptation (Yu et al., 2024).

However, despite their explanatory value, dominant ecosystem perspectives remain limited in at least two important respects. First, they disproportionately privilege urban-industrial contexts characterized by formal institutional arrangements, advanced infrastructure, and technologically mature innovation systems. Second, they tend to conceptualize knowledge primarily in codified or organizational terms, underrepresenting tacit, experiential, and culturally embedded forms of knowledge that are central in traditional productive communities.

These limitations are especially relevant in artisan production systems, where innovation frequently emerges through informal experimentation, collective adaptation, and intergenerational knowledge transmission rather

than through conventional R&D structures. As a consequence, mainstream ecosystem theory remains insufficiently equipped to explain how innovation emerges in territorially embedded contexts where cultural knowledge functions not merely as heritage, but as an active, productive resource.

2.2 Dual transition and contextual sustainability

The dual transition framework conceptualizes digital transformation and environmental sustainability as mutually reinforcing pathways for contemporary economic and territorial development (Kovacic et al., 2024; Quirico & Baber, 2024; Døblen, 2023). This perspective reflects the increasing recognition that sustainable transformation cannot be pursued independently of technological adaptation, nor can digital modernization be considered neutral in environmental terms. Instead, digital technologies are increasingly positioned as strategic enablers of eco-innovation, environmental monitoring, process optimization, and sustainable organizational transformation (Chen et al., 2023; El Samad, 2024; Indiran et al., 2025; Tsouri, 2022). Nevertheless, the implementation of dual transition trajectories remains highly uneven across territorial contexts. While technologically mature industrial environments may rapidly integrate advanced digital infrastructures, rural and culturally embedded production systems often face structural constraints, including limited technological access, reduced institutional support, fragmented innovation capabilities, and dependency on traditional production practices.

This unevenness suggests that dual transition should not be interpreted as a universal technological diffusion pathway. Rather, it must be understood as a context-sensitive socio-technical transformation process in which sustainability outcomes depend on how external technologies are adapted to local cultural, productive, and institutional realities. In artisan communities, this contextual dimension becomes particularly important because productive sustainability is deeply intertwined with inherited tacit knowledge, localized experimentation, and socio-cultural identity. Therefore, dual transition in such settings cannot be reduced to technological modernization alone; it requires adaptive mechanisms capable of integrating digital tools with culturally embedded productive intelligence.

2.3 Knowledge recombination and collective intelligence

Knowledge recombination provides a critical explanatory mechanism for understanding innovation in heterogeneous and territorially embedded environments. Rather than conceptualizing innovation as the production of entirely new knowledge, recombination theory emphasizes the strategic reconfiguration of existing knowledge domains into novel combinations capable of generating adaptive solutions and innovation outcomes (Xiao et al., 2022; Luo, 2022). In artisan production contexts, this process acquires distinctive characteristics because innovation does not primarily emerge through formal scientific research or structured technological development, but through interactions between tacit artisanal expertise, practical experimentation, academic mediation, and external technical knowledge. The integration of these domains creates hybrid innovation pathways that differ significantly from conventional firm-based innovation models.

Recent scholarship has also emphasized that the structure and dynamics of knowledge networks significantly influence innovation performance, particularly where heterogeneous relationships facilitate distributed creativity and adaptive recombination (Yu et al., 2024). In this sense, innovation depends not only on the existence of diverse knowledge resources but also on the capacity of actor networks to articulate, interpret, and transform them collaboratively. This perspective aligns with broader notions of collective intelligence, where distributed actors jointly generate problem-solving capacity through interaction, shared interpretation, experimentation, and contextual adaptation (Yang, 2024). Innovation in territorially embedded systems should therefore be understood not as an isolated invention, but as an emergent socio-cognitive process arising from collective knowledge interaction (see Figure 1).

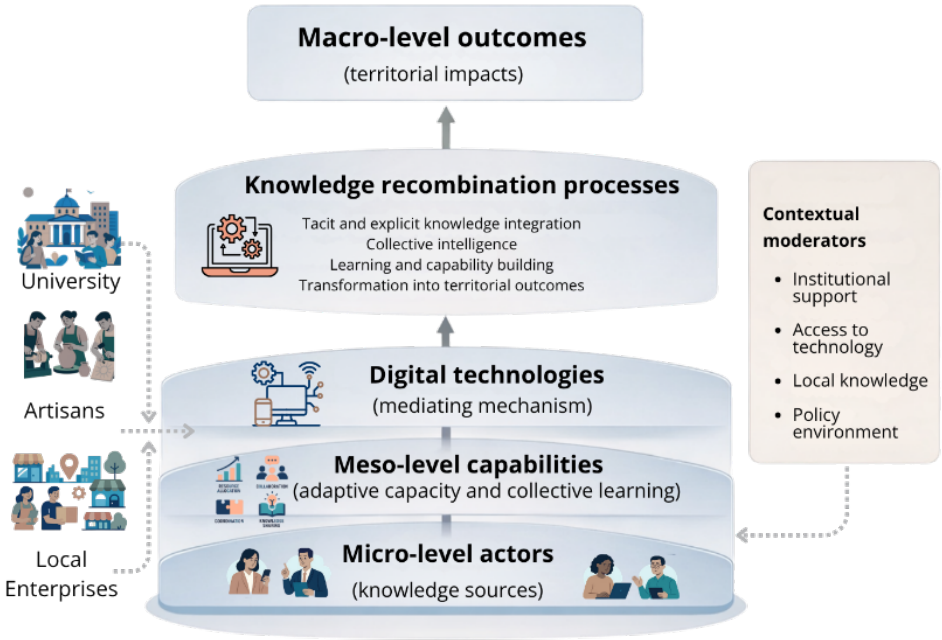


Figure 1. Conceptual model of Intelligent Knowledge Ecosystems that integrates actors, processes, technologies and results. Source: Authors.

2.4 Intelligent Knowledge Ecosystems (IKEs)

In response to these theoretical limitations, this study proposes Intelligent Knowledge Ecosystems (IKEs) as a novel conceptual framework. IKEs are defined as:

territorially embedded socio-technical systems in which collective intelligence, knowledge recombination, and accessible digital-environmental capabilities interact to enable adaptive, inclusive, and sustainable innovation.

As illustrated in Figure 1, the IKE framework integrates four core dimensions:

- Actors.
- knowledge processes.
- enabling technologies.
- innovation outcomes.

Unlike traditional innovation ecosystems, which emphasize institutional structure and economic exchange, IKEs explicitly incorporate tacit community knowledge, cultural identity, and contextual adaptation. Technology within IKEs is not conceptualized as an external determinant of change, but as an enabler of collective learning, experimentation, and knowledge articulation. Similarly, unlike knowledge networks focused on information circulation, IKEs emphasize transformation through recombination and contextual application. This conceptual distinction positions IKEs as an analytically distinct construct designed to explain innovation dynamics in underrepresented territorial contexts.

3. Methodology

3.1 Research design

This study adopts a Participatory Action Research (PAR) design, selected because of its suitability for contexts where knowledge generation is inseparable from intervention, collaboration, and social transformation. Unlike conventional observational methodologies, PAR enables researchers to actively engage with participants in iterative cycles of diagnosis, experimentation, implementation, and reflection, thereby making it particularly appropriate for analyzing emergent innovation dynamics in territorially embedded socio-technical systems. The methodological rationale also responds directly to the conceptual foundations of Intelligent Knowledge Ecosystems (IKEs). Because the proposed framework emphasizes collective intelligence, knowledge recombination, and contextual adaptation, a passive observational design would be insufficient to capture the mechanisms through which innovation emerges. Instead, PAR enables the activation and observation of knowledge interactions in real time.

As illustrated in Figure 2, the methodological design followed a cyclical structure in which intervention and analysis evolved simultaneously. This design allowed the research not only to document outcomes but also to trace the causal mechanisms through which ecosystem dynamics emerged. The intervention unfolded across four iterative phases.

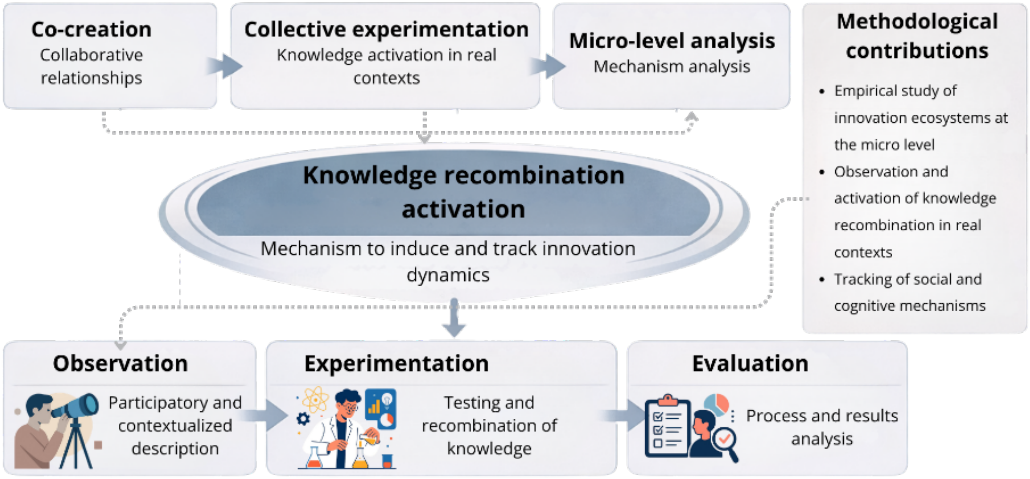


Figure 2. Methodological design based on participatory action research for the activation of knowledge recombination in innovation ecosystems at the micro level. Source: Authors.

1. The first phase involved diagnostic exploration, during which researchers engaged with local actors to identify productive challenges, sustainability constraints, and existing knowledge practices. This stage established the contextual baseline and clarified ecosystem conditions.
2. The second phase focused on collaborative planning, where artisans, students, and academic facilitators co-designed intervention strategies, emphasizing practical experimentation and mutual learning.
3. The third phase consisted of implementation and co-creation, during which workshops, technological experimentation, and sustainability-oriented interventions were conducted.
4. The fourth phase involved reflection, interpretation, and ecosystem consolidation, allowing participants to evaluate outcomes, refine practices, and collectively reinterpret the innovation process.

This cyclical structure strengthened methodological coherence while addressing reviewer concerns regarding insufficient explanation of research procedures.

3.2 Case study

The empirical setting for this study was the project "Tradición y Cultura Viva: Energía Creativa para la Sustentabilidad en la Artesanía de Guanajuato", implemented in the Lindavista neighborhood of Dolores Hidalgo, Guanajuato, Mexico. This setting represents a highly relevant case for studying dual transition in underrepresented territorial contexts. Dolores Hidalgo possesses a strong artisan identity centered on ceramic production, where productive practices rely heavily on inherited tacit knowledge, experiential learning, and culturally embedded techniques. At the same time, the community faces structural constraints related to technological access, productive competitiveness, environmental sustainability, and market adaptation.

These characteristics make the setting analytically valuable because it combines high cultural knowledge intensity with low formal technological sophistication. Rather than functioning merely as a passive context, the territory itself constituted an active component of the ecosystem, shaping interaction patterns, innovation constraints, and adaptation opportunities.

3.3 Participants

The study involved 48 participants, selected through purposive sampling to maximize heterogeneity of knowledge sources and ecosystem interaction. Participants included 22 artisans, whose expertise in ceramic production represented the tacit and culturally embedded knowledge base of the ecosystem. The study also involved 18 university students from the Industrial/Business Management Engineering program, who acted as technological intermediaries, learners, and co-creators.

Additionally, 8 researchers and institutional actors participated, including academic staff, project facilitators, and local stakeholders responsible for ecosystem articulation. This composition was methodologically intentional. Rather than treating participants as isolated respondents, the research positioned them as ecosystem actors whose interactions generated the dynamics under analysis.

The inclusion of heterogeneous participants strengthened the possibility of observing recombination processes between tacit, academic, and technical knowledge domains.

3.4 Data collection

Data collection was conducted over a 12-month intervention period, enabling longitudinal observation of ecosystem evolution and innovation dynamics. To strengthen empirical robustness, the study employed methodological triangulation, the study employed methodological triangulation, combining multiple qualitative sources and interaction settings.

First, 25 semi-structured interviews were conducted with artisans, students, and institutional actors. These interviews explored perceptions regarding innovation, sustainability, technological adaptation, and collaborative learning.

Second, six co-creation workshops were implemented as intervention spaces where ecosystem actors collaboratively developed and tested solutions. These workshops served simultaneously as empirical sites of innovation activation and observation.

Third, field observations were conducted throughout productive activities and collaborative sessions. These observations allowed researchers to capture interaction patterns, adaptive practices, and real-time decision-making dynamics.

This triangulated strategy increased empirical depth while enabling both outcome documentation and mechanism tracing. of observing recombination processes between tacit, academic, and technical knowledge domains.

3.5 Analytical strategy and rigor

Data were analyzed using a qualitative process-oriented approach combining thematic coding and process tracing. The analysis proceeded in three stages:

The first stage involved open coding, aimed at identifying emergent categories associated with knowledge recombination, collaboration, sustainability practices, and technological mediation.

The second stage involved axial coding, where relationships between categories were examined to identify ecosystem patterns and causal connections.

The third stage employed process tracing, allowing the reconstruction of event sequences through which innovation dynamics emerged.

This analytical strategy was particularly important because the objective was not simply descriptive documentation, but explanation of underlying causal mechanisms. To enhance rigor and trustworthiness, several validation strategies were applied. First, data triangulation was achieved through multiple empirical sources. Second, source triangulation was ensured through the participation of heterogeneous actor groups. Third, iterative member reflection was incorporated through ongoing interaction with participants during the action research process. These procedures strengthened interpretive validity and methodological rigor (see Table 1).

Table 1: Data collection and analysis framework.

Method	Description	Participants	Purpose
Interviews	Semi-structured	25	Identify knowledge processes
Workshops	Co-creation sessions	48	Activation and collaborative learning
Observation	Fieldwork	All	Contextual Understanding

4 Results

4.1 Knowledge recombination as the central innovation mechanism

The findings indicate that innovation within the ecosystem did not emerge through isolated technological adoption, but through structured processes of knowledge recombination. Consistent with recombination theory, innovation was generated through the integration of previously distinct knowledge domains (Xiao et al.,

2022). However, unlike formal innovation environments, the recombination observed here involved tacit artisanal expertise interacting with technical and academic knowledge.

Artisans contributed a deep experiential understanding of ceramic materials, firing behavior, pigment preparation, and production routines.

Students and researchers introduced external knowledge related to monitoring, experimentation, documentation, and process optimization.

The interaction between these domains produced adaptive innovations rather than direct technological substitution. This finding strengthens the theoretical argument that innovation in territorially embedded systems emerges through socio-cognitive recombination rather than technological determinism.

4.2 Technology as a learning mediator rather than an external driver

One of the most significant findings concerns the role of technology. Contrary to conventional narratives in dual transition literature, technology did not operate as an autonomous transformation engine. Instead, it functioned as a mediator of collective learning. Digital and accessible tools were introduced through participatory processes, allowing actors to gradually integrate them into existing productive routines. Students played an intermediary role by translating technical knowledge into accessible applications, while artisans adapted these tools to their practical realities. This resulted in forms of technological appropriation rather than passive adoption. Technologies supported:

- Temperature monitoring.
- Production documentation.
- Process organization.

More importantly, these tools made previously implicit production processes visible, enabling reflection, experimentation, and collaborative decision-making. Thus, digitalization acted less as infrastructure and more as cognitive mediation within the ecosystem.

4.3 Sustainability and productive efficiency

The intervention also generated measurable sustainability-oriented improvements. Participants reported greater productive consistency through improved firing process control, reduced material waste, and more efficient energy use. Adjustments in kiln loading, thermal monitoring, and production sequencing contributed to lower operational inefficiencies. These outcomes suggest that sustainability improvements emerged not from advanced technological infrastructure, but from context-sensitive adaptation and learning. This finding directly addresses reviewer concerns regarding whether the results represent empirical observations or merely expected outcomes. The evidence indicates that sustainability gains were observable consequences of ecosystem interaction.

4.4 Cultural identity as a strengthened rather than threatened dimension

A particularly important finding concerns cultural identity; a common assumption in modernization discourse is that technological intervention erodes traditional practices. However, the findings suggest the opposite.

Participants reported that innovation processes reinforced appreciation for local knowledge, encouraged reinterpretation of traditional techniques, and enhanced product differentiation.

Rather than replacing artisanal identity, innovation increased its perceived value.

This suggests that dual transition in culturally embedded systems need not produce homogenization, but may instead strengthen socio-cultural resilience.

4.5 Ecosystem emergence through collective intelligence

Across the intervention, innovation increasingly shifted from isolated problem-solving toward distributed collective intelligence. Knowledge sharing became more intentional, experimentation more collaborative, and decision-making more distributed. These patterns indicate the emergence of ecosystem properties rather than isolated collaborative episodes. The findings therefore empirically support the IKE framework proposed in Figure 1, demonstrating how actors, knowledge processes, enabling tools, and sustainability-oriented outcomes interact dynamically.

5 Discussion

The findings of this study provide important theoretical and practical insights into how innovation emerges in territorially embedded production systems that have traditionally remained underrepresented in mainstream ecosystem scholarship. The first major contribution concerns the reconceptualization of innovation ecosystems. Conventional literature has predominantly framed innovation ecosystems as formal interorganizational configurations driven by firms, institutions, and market-oriented exchange (Oliveira et al., 2025; Pattinson et al., 2023). However, the empirical evidence presented here demonstrates that innovation may emerge under fundamentally different structural conditions. In the case examined, innovation did not depend on formal R&D infrastructure, advanced technological sophistication, or institutional complexity. Instead, it emerged through socially mediated processes of interaction, collective experimentation, and contextual adaptation among heterogeneous actors.

These findings extend ecosystem scholarship by demonstrating that innovation ecosystems are not inherently restricted to technologically mature industrial environments. Rather, ecosystem dynamics may emerge wherever diverse actors interact through mechanisms capable of enabling knowledge transformation, adaptive experimentation, and collaborative problem-solving. From this perspective, the concept of Intelligent Knowledge Ecosystems (IKEs) is not proposed as a semantic variation of conventional innovation ecosystems, but as an analytically distinct construct designed to explain innovation dynamics in contexts where tacit, cultural, and territorially embedded knowledge constitute the primary innovation resource.

The second major contribution concerns the explanatory role of knowledge recombination. Existing scholarship has identified recombination as a central mechanism for innovation generation, particularly in technologically intensive or formally structured environments (Xiao et al., 2022; Luo, 2022). However, empirical evidence from rural and artisan production systems remains comparatively limited. The present findings demonstrate that recombination operates effectively even in environments characterized by low formal technological maturity. More importantly, the nature of recombination differs significantly from dominant innovation models. Rather than combining codified scientific or organizational knowledge, the ecosystem analyzed here recombined tacit artisanal expertise, experiential production knowledge, academic mediation, and accessible technological inputs. Innovation therefore, emerged not through technological substitution but through socio-cognitive hybridization. This expands recombination theory into underrepresented territorial contexts and suggests that innovation capacity may depend less on technological intensity than on the architecture of interaction through which heterogeneous knowledge domains are articulated.

A third theoretical contribution relates to dual transition scholarship. Much of the existing literature conceptualizes dual transition primarily through macro-level perspectives focused on policy, infrastructure, and institutional transformation (Kovacic et al., 2024; Quirico & Baber, 2024). While these perspectives offer valuable insights, they often understate how digital and environmental transitions materialize at micro-territorial scales. The findings presented here indicate that dual transition may emerge through localized participatory learning processes rather than exclusively through top-down infrastructure deployment.

Technology played a fundamentally different role from that assumed in many deterministic transition narratives. Rather than functioning as an autonomous transformation engine, technology operated as a mediator of collective learning, knowledge articulation, and adaptive experimentation. This shifts the conceptual interpretation of dual transition away from technological determinism toward a socio-technical adaptation perspective grounded in collaborative learning. Another important contribution concerns cultural sustainability. Mainstream modernization narratives frequently position innovation and tradition as opposing forces, implying that technological intervention necessarily weakens local identity or traditional productive knowledge. However, the empirical evidence suggests the opposite. Artisans did not perceive technological mediation as threatening authenticity. Instead, the innovation process increased awareness of the strategic value of local knowledge and strengthened confidence in culturally grounded productive practices. This finding expands sustainability debates by incorporating sociocultural resilience alongside environmental and economic performance.

Finally, the study contributes to understanding the role of universities in ecosystem formation. Higher education institutions are often framed as knowledge producers, technology transfer actors, or innovation intermediaries. However, the findings suggest a more active and structurally significant role. In the present case, the university functioned not merely as an external provider of expertise but as an ecosystem orchestrator that enabled interaction between heterogeneous knowledge domains, facilitated collaborative experimentation, and created the socio-technical conditions required for adaptive innovation. This interpretation aligns with emerging perspectives from triple helix and university ecosystem scholarship, which increasingly recognize universities as strategic actors in sustainability-oriented innovation systems (Secundo et al., 2024; Shkarupeta et al., 2024).

Rather than transferring knowledge unidirectionally, the university contributed to collective learning, knowledge translation, and ecosystem stabilization. This expands conventional interpretations of university engagement by positioning higher education institutions as catalytic actors within territorially embedded dual transition processes.

6 Conclusions

This study examined how innovation emerges in artisan production contexts under the framework of the dual transition and proposed the concept of Intelligent Knowledge Ecosystems (IKEs) as a new analytical lens for understanding these dynamics. The findings demonstrate that innovation in territorially embedded communities does not depend primarily on formal institutional complexity or advanced technological infrastructure. Instead, it emerges through structured processes of collective intelligence, knowledge recombination, and contextual technological mediation. This insight represents the central contribution of the study; By introducing IKEs, the research extends conventional innovation ecosystem theory toward culturally grounded and territorially sensitive environments traditionally underrepresented in mainstream scholarship. The empirical findings validate this framework by showing how heterogeneous actors collaboratively generate innovation capacity through adaptive interaction. The study also contributes to dual transition scholarship by demonstrating that digital and environmental transformation can emerge from participatory ecosystem processes at the micro-territorial level. This challenges deterministic narratives that equate transformation with technological sophistication.

Equally important, the findings reveal that innovation and cultural preservation are not inherently contradictory. On the contrary, innovation processes may reinforce cultural identity when technological mediation is designed around contextual adaptation and collective ownership. In practical terms, the study provides a transferable model for universities, communities, and policy actors seeking more inclusive innovation pathways. Ultimately, Intelligent Knowledge Ecosystems offer a promising framework for rethinking sustainable territorial innovation in contexts where knowledge is cultural, distributed, and deeply embedded in local practice.

7 Acknowledgments

This publication is part of the Scientific and Humanistic Research Projects in Strategic Axes 2025, Budget Program F003 (PpF003) "National Strategic Programs of Science, Technology and Linkage with the Social, Public and Private Sectors" Project PEE-2025-C-6 of the Secretariat of Science, Humanities, Technology and Innovation SECIHTI. / This publication is part of the Scientific and Humanistic Research Projects in Strategic Axes 2025, Budget Program F003 (PpF003) "National Strategic Programs of Science, Technology and Linkage with the Social, Public and Private Sectors" Project PEE-2025-C-6 of the Secretariat of Science, Humanities, Technology and Innovation SECIHTI.

8 Ethics Declaration

This study was conducted in accordance with the ethical principles established for academic research. It did not involve the participation of humans or animals. In case of data use, these were obtained from public sources or processed anonymously, guaranteeing confidentiality and compliance with current regulations.

9 AI Declaration

The authors state that artificial intelligence tools were used to support the writing, organization, and revision of the text. However, the research design, data analysis, and conclusions were developed entirely by the authors, who assume full responsibility for the content.

10 Conflict of Interest Statement

The authors declare no conflicts of interest in relation to the publication of this article.

References

Burtscher, J., Leipziger, M., Kanbach, D. K., & Kraus, S. (2025). Pathways to twin transformation in SMEs: The role of innovation ecosystems. European Journal of Innovation Management, 29(2), 464–489. https://doi.org/10.1108/EJIM-11-2024-1382

Chen, X. (2023). Enabling the twin transitions: Digital technologies support environmental sustainability through lean principles. *Sustainable Production and Consumption*, *38*, 13–27. https://doi.org/10.1016/j.spc.2023.03.020

Costa, J., & Moreira, A. C. (2022). Public policies, open innovation ecosystems and innovation performance: Analysis of the impact of funding and regulations. *Journal of Open Innovation: Technology, Market, and Complexity*, *8*(4), 210. https://doi.org/10.3390/joitmc8040210

Dæhlen, M. (2023). *The twin transition century: The role of digital research for a successful green transition of society?* The Guild Insight Paper No. 5. https://doi.org/10.48350/184458

El Samad, S. (2024). Eco-innovation: Transforming into a sustainable future. *Proceedings*, *101*(1), 14. https://doi.org/10.3390/proceedings2024101014

Heaton, S., & Min, J. (2025). Open innovation in ecosystems: Exploring how the affiliation of an ecosystem partner impacts the benefits of collaboration. *Research Policy*, *54*(2), 105168. https://doi.org/10.1016/j.respol.2024.105168

Indiran, L., Fu, C., & Fahim, N. A. (2025). Eco-innovation methodologies: A literature review. *Discover Sustainability*, *6*, 1143. https://doi.org/10.1007/s43621-025-01621-y

Kovacic, Z., García Casañas, C., & Argüelles, L. (2024). The twin green and digital transition: High-level policy or science fiction? *Environment and Planning E: Nature and Space*, *7*(6). https://doi.org/10.1177/25148486241258046

Luo, Z. (2022). Knowledge recombination effects on innovation quality: Evidence from the pharmaceutical industry. *European Journal of Innovation Management*. https://doi.org/10.1108/EJIM-10-2021-0517

Oliveira, J. P., Figueiredo, M. D., & Teixeira, A. A. C. (2025). Innovation ecosystems and open innovation in micro-enterprises. *Journal of Innovation & Knowledge*. https://www.sciencedirect.com/science/article/pii/S2199853124002373

Pattinson, S., Nicholson, J. D., Ehret, M., Velu, C., & Ryan, P. (2023). Innovation ecosystems in B2B contexts: Owning the space. *Industrial Marketing Management*, *108*, A1–A9. https://doi.org/10.1016/j.indmarman.2022.11.004

Quirico, O., & Baber, W. F. (2024). Twin transitions? Implementing climate policies in the European Union through digital transformation. En O. Quirico & W. F. Baber (Eds.), *Implementing climate change policy: Designing and deploying net zero carbon governance*. Cambridge University Press. https://doi.org/10.1017/9781009341493.012

Remneland Wikhamn, B., & Styhre, A. (2023). Open innovation ecosystem organizing from a process view: A longitudinal study in the making of an innovation hub. *R&D Management*, *53*(1), 24–42. https://doi.org/10.1111/radm.12537

Shkarupeta, E., & Babkin, A. (2024). Eco-innovative development of industrial ecosystems based on the quintuple helix. *International Journal of Innovation Studies*, *8*(3), 273–286. https://doi.org/10.1016/j.ijis.2024.04.002

Tsouri, M. (2022). Knowledge recombination for emerging technological fields: The case of green shipping. *Technovation*, *114*, 102460. https://www.sciencedirect.com/science/article/pii/S0166497222000013

Xiao, Y., Wang, C., van Knippenberg, D., & Li, N. (2022). A knowledge recombination perspective of innovation: Review and future research agenda. *Journal of Management*, *48*(7), 1765–1797. https://doi.org/10.1177/01492063211055982

Xue, X., Dhumras, H., Thakur, G., & Shukla, V. (2025). Integrating artificial intelligence and sustainable materials for smart eco-innovation in production. *Scientific Reports*, *15*, 36942. https://doi.org/10.1038/s41598-025-20803-2

Yang, G. (2024). Knowledge element relationship and value co-creation in the innovation ecosystem. *Sustainability*, *16*(10), 4273. https://doi.org/10.3390/su16104273

Ying, Y., & Jin, S. (2024). Artificial intelligence and green product innovation: Moderating effect of organizational capital. *Heliyon*, *10*(7), e28572. https://doi.org/10.1016/j.heliyon.2024.e28572

Yu, L., Wang, X., Li, J., & Peng, M. W. (2024). The role of relation dynamics and knowledge network decomposability in firm creativity and innovation. *Science & Public Policy*, *51*(4), 602–615. https://doi.org/10.1093/scipol/scad048

The New Organizational Knowledge Dynamics Model that Integrates Human and Artificial Knowledge

Constantin Bratianu[1], Ruxandra Bejinaru[2]
[1]Bucharest University of Economic Studies, Bucharest, Romania
[2]University "Stefan cel Mare" of Suceava, Suceava, Romania

Abstract

The emergence of Generative Artificial Intelligence (GenAI) and its explosive applications challenges the structure, functionality and efficiency of the Knowledge Management Systems (KMS). GenAI processes data, information, and human knowledge and generates artificial knowledge that belongs to the sciences of the artificial. KMS have been designed to create, share, transfer, and use human knowledge. The basic theory of organizational knowledge creation dynamics has been developed by Ikujiro Nonaka and Hirotaka Takeuchi and illustrated as the well-known SECI (Socialization, Externalization, Combination, Internalization) model. Knowledge is created through a series of transformations between tacit knowledge and explicit knowledge, and scaling up from individuals to teams and organizations. GenAI transformed KMS into hybrid systems composed of people and machines. The problem now is how to integrate human knowledge with artificial knowledge to enhance the efficiency of the KMS. From a practical viewpoint, the problem is how to augment the SECI model with artificial knowledge such that it can represent the new Hybrid Knowledge Management Systems (HKMS). The difficulty comes from the fact that human knowledge and artificial knowledge are two distinct constructs, and they cannot be mixed entirely. Human knowledge is a synthesis of rational, emotional, and spiritual knowledge, while artificial knowledge is exclusively rational. Therefore, the augmentation of each stage of the SECI model with artificial knowledge cannot be a solution because tacit or emotional knowledge is embodied. The models published so far focused on this solution, ignoring the fact that tacit knowledge cannot be integrated directly with artificial knowledge. The purpose of this research is to show a new perspective on augmenting the SECI model with artificial knowledge. We consider a separate cycle for artificial knowledge that mirrors the structure of the SECI cycle, and an interaction between them through the process of combination (i.e., the only process that is based exclusively on explicit knowledge). Therefore, the new complex model is composed of two cycles (i.e., the human knowledge and the artificial knowledge cycles), which are interconnected through the combination process.

Keywords: human knowledge, artificial knowledge, GenAI, SECI model, knowledge creation dynamics

Paper type: Academic Research Paper

1 Introduction

The theory of knowledge creation dynamics within Knowledge Management Systems (KMS) was initiated by Nonaka (1994) and further developed by Nonaka and Takeuchi (1995, 2019), Nonaka, Toyama, and Hirata (2008), and Konno (2024). This theory reflects the cyclic transformations between tacit knowledge and explicit knowledge on the epistemological axis, and the upscaling of knowledge integration from individuals to teams and organizations along the ontological axis. Based on these theories Nonaka and his co-authors developed the well-known SECI (Socialization, Externalization, Combination, Internalization) model. Socialization is the process of tacit knowledge sharing between people by seeing and imitating. It is a knowledge sharing using non-verbal communication because tacit knowledge is an embodied knowledge (Davenport & Prusak, 2000; Hadjimichael & Tsoukas, 2019; Polanyi, 1983; Tsoukas, 2011). Socialization is performed within a social context. Externalization is the process of transformation of tacit knowledge into explicit knowledge in the brain of a certain individual. This transformation is achieved through natural or symbolic language, analogies, and metaphors (Lakoff & Johnson, 1999; Pinker, 2008). Therefore, explicit knowledge can be transferred between people as a result of using natural language. We may say that explicit knowledge is codified knowledge. Combination is the process of explicit knowledge sharing within a social context. Ideas expressed by individuals are integrated, amplified and refined until a final agreement is achieved. Combination is based exclusively on explicit or rational knowledge. It is the most powerful process for knowledge co-creation and upscaling along the ontological axis. Internalization is the process of transformation of explicit knowledge into tacit knowledge at the individual level. It contributes to a better understanding of the events and phenomena as a result of learning. The whole knowledge cycle develops in time, generating a knowledge spiral (Nonaka & Takeuchi, 2019). This time dimension is a third dimension of the whole knowledge creation process.

The emergence of artificial knowledge that is generated through the interaction between people and Generative Artificial Intelligence (GenAI) requires an augmentation of the SECI model to integrate it. However,

artificial knowledge is different from human knowledge, although its roots are human data, information and knowledge (Baker, 2023, 2025; Bratianu, 2025; Di Vaio et al., 2024; Kurzweil, 2024; Lee & Qiufan, 2024). While human knowledge is a synthesis of rational, emotional, and spiritual knowledge (Bratianu & Bejinaru, 2019, 2020), artificial knowledge is exclusively rational and explicit (Russell & Norvig, 2023). Human knowledge is created within a semantic field, while artificial knowledge is generated by GenAI based on syntactic rules and statistical distributions of words within a given database. Computers do not think. They make algorithms to work and generate artificial knowledge. "Current GenAI models don't think or create things *per se*, but instead generate new things from parts of old things found in its database. (The term "things" in this context being images, videos, numbers, or text, depending on the GenAI application you are using). A GenAI output is the model's best prediction of what you are seeking" (Baker, 2025, p. 11). Artificial knowledge is generated by a series of mathematical algorithms that process a huge database. If the input data contains errors, the resulting artificial knowledge will incorporate those errors. Users should never trust the content of artificial knowledge. Instead, they should check it all the time against some evidence. While human knowledge represents a *justified true belief* (Audi, 2011; Nonaka & Takeuchi, 1995), artificial knowledge does not reflect such a true belief. All of these arguments demonstrate that integrating human knowledge with artificial knowledge implies a new perspective of organizational knowledge creation dynamics that creates a knowledge gap in the literature. As a consequence, we formulate the following research question:

> *RQ: How to augment the SECI model with artificial knowledge within the hybrid knowledge management systems in knowledge-intensive organizations?*

The structure of this paper is as follows: after this short introduction, we perform a critical literature review, then we show the method used for the present research, followed by the results and discussions. Finally, there is a section of concluding remarks and research limitations.

2 Literature Review

The purpose of our critical literature review is to analyse some proposed models for augmenting the SECI model with artificial knowledge as a result of the interaction between people and GenAI applications within the hybrid knowledge management systems. Jarrahi et al. (2023) analyse the partnership between AI and knowledge management, considering knowledge creation, knowledge storing and retrieving, knowledge sharing, and knowledge application. Knowledge creation may occasionally start from scratch. More frequently, it is a process of restructuring and recombination of the known facts within a new perspective. GenAI cannot think and cannot develop experiential knowledge (Kolb, 2015). However, it has the capability of performing a complex statistical analysis of huge databases and identifying unknown correlations between them. Based on them, GenAI creates new patterns for the prediction of new evolutions. "As a result of this capacity to derive their own rules based on similar patterns in the data, deep-learning AI can discover patterns in the available data sets previously unknown to the organization" (Jarrahi et al., 2023, p. 88). GenAI is very efficient in searching for data, information and knowledge integrated in the huge databases, classifying, organizing, storing and retrieving explicit knowledge. Also, it is of great help in amplifying knowledge sharing by connecting people with similar knowledge interests, regardless of the space and time separating them. The authors introduce the concept of *human-AI symbiosis* that reflects a new aspect of the interaction between people and AI that leads to achieving collaborative intelligence. However, they stress the fact that this process refers only to explicit knowledge. Tacit knowledge remains an attribute of the human agents.

Harfouche et al. (2022) developed a recursive theory of knowledge augmentation by integrating human intuition and knowledge in artificial intelligence to augment organizational knowledge. The authors focus on the human-GenAI interaction and design a new framework for organizational knowledge creation by augmenting the SECI model. "The Recursive Theory of Knowledge Augmentation (the KAM model) explains how tacit and explicit knowledge held by individuals and organizations can be simultaneously enlarged and enriched through the recursive process enabled by human-in-loop Informed AI" (Harfouche et al., 2022, p. 2). The approach aims at integrating human knowledge into the data-driven machine learning process, having the SECI model as a basic framework. From a practical viewpoint, the authors transformed the 2X2 matrix of the tacit-explicit knowledge transformations contained in the SECI model into a 3X3 matrix of the tacit-explicit-artificial knowledge transformations. Therefore, the new augmented matrix contains the SECI knowledge transformations (i.e., socialization, externalization, combination, and internalization) and five new processes that include artificial knowledge. *Alimentation* is the transformation of tacit knowledge into artificial knowledge. It is about how human intuition feeds the training of the Machine Learning. *Aggregation* is the transformation of explicit knowledge into artificial knowledge. The authors show that the human-in-the-loop Informed AI (IAI) was trained using explicit object-based knowledge. *Amplification* is the reverse process of the transformation of artificial

knowledge into explicit knowledge. The human-in-the-loop IAI can amplify experts' explicit knowledge. *Reflection* is the transformation of artificial knowledge into tacit knowledge. The authors consider that new artificial knowledge can contribute to human reflection and tacit knowledge creation. The last new process is called *fusion*, and it refers to the transformation of artificial knowledge into another artificial knowledge. "It results from the recursive loops needed to design, train, test, and implement a human-in-the-loop IAI in an organization. The outputs of the first loop of Alimentation, Aggregation, Amplification, and Reflection can set off or trigger a second loop" (Harfouche et al., 2022, p. 10). The analysis of the new framework shows that tacit knowledge has direct interaction with artificial knowledge through alimentation and reflection, although tacit knowledge is embodied (Hadjimichael & Tsoukas, 2019; Nonaka & Takeuchi, 1905; Tsoukas, 2011) and cannot have such a direct interaction. This is a weakness of the new framework and of the idea that the SECI model can be augmented through mixing human and artificial knowledge.

Böhm and Durst (2025) were driven by the same idea of augmenting the SECI model with artificial knowledge, but they sought another approach. They kept the structure of the SECI model and proposed a mix between human knowledge and artificial knowledge within each stage of the model. The result is the GRAI – Generative Receptive Artificial Intelligence model. The basic idea of the GRAI model is to consider each stage of the SECI model and propose its extension by integrating artificial knowledge. For instance, for socialization, the authors consider that "virtual environments or platforms could be created where people in different (geographical) locations can interact in real time, fostering socialization regardless of physical distance. These platforms can simulate face-to-face interactions, enabling spontaneous exchanges of tacit knowledge" (Böhm & Durst, 2025, p. 4). The authors face the problem like Harfouche et al. (2022): misunderstanding of tacit knowledge and the mechanism of socialization designed by Nonaka and Takeuchi (1995, 2019) in SECI. People can work with virtual environments and digital platforms using their explicit knowledge, not tacit knowledge. In the vision of Nonaka and Takeuchi (1995, 2019), socialization means the exchange of tacit knowledge through non-verbal communication. Digital platforms imply first transforming tacit knowledge into explicit knowledge using natural or symbolic languages, and only then performing any exchange of that knowledge. The same problem appears when considering externalization and internalization due to tacit knowledge. The only process open to mixing human and artificial knowledge remains combination. Therefore, the GRAI model designed for SECI model augmentation with artificial knowledge does not work. It is necessary to consider some other approaches.

The purpose of this paper is to show that this barrier of tacit knowledge can be overcome if we let the SECI model be as it is and consider a new cycle for artificial knowledge. The new cycle can be connected with SECI model through the combination stage because it is the only process involving exclusively explicit knowledge.

3 Methodology

This is a conceptual paper aiming at developing a new model for knowledge creation within hybrid knowledge management systems in knowledge-intensive organizations. It is based on a critical literature review, the theory of knowledge fields and knowledge dynamics (Bratianu & Bejinaru, 2019, 2020), the theory of knowledge creation dynamics (Nonaka & Takeuchi, 1995, 2019), and the design science research (Gregor & Hevner, 2013; Hevner et al., 2004). The theory of knowledge fields and knowledge dynamics states that we should understand knowledge as a complex integration of three basic fields: rational knowledge, emotional knowledge, and spiritual knowledge. The theory of knowledge dynamics states that a field of knowledge can be transformed into another field of knowledge (e.g., rational knowledge can be transformed into emotional knowledge or spiritual knowledge). The theory of knowledge creation developed by Nonaka and Takeuchi (1995, 2019) is incorporated in the SECI model that has been described above. The theory of design science research shows how to create artefacts that offer solutions to identified problems in information science, engineering, and management. Those artefacts can be new concepts, ideas, models, or theories within the specificity of the sciences of the artificial (Simon, 1996). According to Henvner et al., "Design science addresses research through the building and evaluation of artefacts designed to meet the identified business need. The goal of behavioural science research is truth. The goal of design science is utility" (Hevner et al., 2004, pp. 79-80).

4 Results and Discussion

As discussed in the previous sections, the proposed augmented models of the human knowledge creation model developed by Nonaka and Takeuchi (1995, 2019), focused on enlarging the SECI model with new processes which are specific to artificial knowledge. This approach ignores the embodied nature of tacit knowledge and creates a false solution. Our approach is to design a new cycle focusing only on artificial knowledge and to connect it with the SECI model through the combination stage, because combination is based on exclusive rational or explicit knowledge. To design a balanced complex model, it is necessary to imagine an

artificial knowledge-generating model that is similar to the human knowledge-creation cycle and that has the capacity to interact with it. We propose a cycle composed of the following stages: combination, artificialization, synthetization, and humanization (CASH) (See Figure 1).

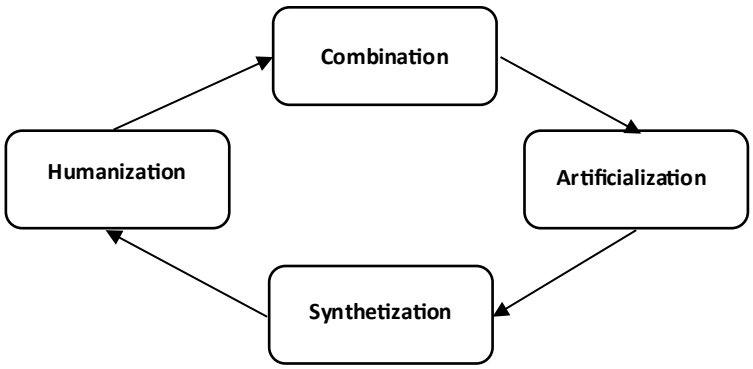


Figure 1: The CASH cycle for the artificial knowledge
(Source: Authors' research)

Combination is the same stage as for SECI model. It is the stage where explicit knowledge from individuals is amplified and integrated with artificial knowledge produced by GenAI. For the CASH cycle, combination means the data selection and its structuring to form the input for the artificial knowledge process. *Artificialization* is the next stage where human data, information, and knowledge are transformed into machine language to be processed. *Synthetization* is the stage where the GenAI algorithms transform the input data into predictive patterns by finding correlations within the huge database used for their training. *Humanization* is the stage where the result existing in the machine language is transformed into humanlike texts, images or videos. It is essential for the interaction between humans and machines. The cycle ends in combination, where it interacts with the SECI model. Therefore, the new model we propose is composed of both these knowledge cycles, SECI and CASH, which interact through the combination stage.

The CASH cycle processes only data, information, and knowledge expressed in a rational or explicit form, overcoming the barrier of dealing directly with tacit knowledge. This is a new perspective that brings value to the proposed CASH cycle and its connection with the SECI cycle. The proposed cycle can be evaluated at this stage using an analytical evaluation (Hevner et al., 2004).

- *Clearly define the model structure*. The CASH model has a similar structure to SECI. It is composed of four stages that contain transformations from the human data, information, and knowledge into artificial data, information, and knowledge and vice versa. Therefore, it is comparable and compatible with the SECI model.
- *Check logical consistency*. The flow of artificial knowledge is continuous and without any barriers because it follows mathematical consistency. Also, the conflict with tacit knowledge has been removed from the cycle.
- *Assess completeness*. The cycle is complete because the final output is the result users need. It is expressed in natural language, which makes the interaction between humans and machines much easier than otherwise.
- *Evaluates parsimony*. The CASH cycle is rather simple and has no redundant elements.
- *Perform ontological analysis*. The cycle starts and ends with the process of combination, which is shared with the human knowledge cycle, and then is performed only by the computer.
- *Theoretical grounds*. The CASH cycle is based on mathematical theories and neural architecture which are specific GenAI applications (Baker, 2023, 2025; Russell & Norvig, 2023).

Figure 2 shows how the SECI cycle interacts with the CASH cycle, forming a new complex framework that is totally different from the GRAI and KAM models discussed in the literature review section.

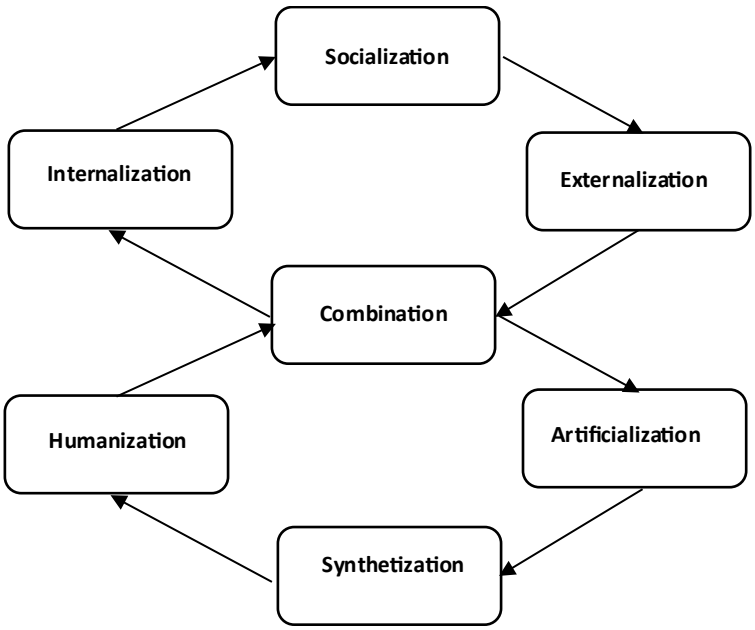


Figure 2: The complex SECI-CASH framework for integration of human and artificial knowledge (Source: Authors' research)

Looking at the new framework, we can see that organizational knowledge creation is a complex process composed of human and artificial knowledge, and the flow of knowledge does not create a spiral anymore. The flow of knowledge suggests a Möbius-strip model, which is more appropriate for the hybrid knowledge management systems than the augmented models proposed so far in the literature (Böhm & Durst, 2025; Harfouche et al., 2022). The new model suggests an interactive and iterative process of knowledge creation that leads to the idea of a strange attractor from the theory of complexity (Bird, 2003).

5 Conclusions

The purpose of this paper is to design a new model for organizational knowledge creation capable of integrating human and artificial knowledge. This new framework is requested by the hybrid knowledge management systems that include GenAI applications. The proposed augmented models of the SECI knowledge cycle focus on considering each of it stage a mixing place for both human and artificial knowledge. Therefore, socialization, externalization, combination, and internalization are augmented with artificial knowledge. However, tacit knowledge is embodied and cannot be mixed together with artificial knowledge that is exclusively rational or explicit. Therefore, augmenting socialization, externalization, and internalization with artificial knowledge makes no sense from the theory of knowledge creation developed by Ikujiro Nonaka and Hirotaka Takeuchi.

The present paper comes with a new perspective on considering a new cycle dedicated to artificial knowledge and being capable of interacting with the human knowledge creation cycle. The interaction can occur at the stage of combination, which contains only explicit knowledge. The new cycle has a similar structure to the SECI cycle and has the capacity to interact with it. It contains the following stages: combination, artificialization, synthetization, and humanization. Combination appears as a shared stage between both cycles because it contains exclusively rational or explicit knowledge. Artificialization is the stage where the human data input is transformed into a machine representation. Synthetization is the stage of mathematical algorithms that realize deep learning through neural networks. Its output represents new patterns based on correlations between the elements of the input data. Humanization is the process of transformation of those patterns expressed in machine language into products expressed in natural language, images or video, which constructs a humanlike dialogue with the users. The cycle is closing or starting a new iteration in the stage of combination. Together, both cycles (i.e., SECI and CASH) constitute a new framework that is capable of generating organizational knowledge for the hybrid knowledge management systems. This new framework distinguishes the present research from those already published (i.e., GRAI and KAM models).

The limitation of the present paper comes from the fact that it is conceptual, without empirical validation. That validation will constitute a further direction of our research.

6 Ethics Declaration

This is a conceptual paper for which there is no need for any ethical clearance.

7 AI Declaration

We have not used AI in preparing this research and writing this paper.

References

Audi, R. (2011). A Contemporary Introduction to the Theory of Knowledge. Third edition. Routledge, London.

Baker, P. (2023). ChatGPT for Dummies. John Wiley & Sons, Hoboken.

Baker, P. (2025). Generative AI for Dummies. John Wiley & Sons, Hoboken.

Bird, R.J. (2003). Chaos and Life: Complexity and Order in Evolution of Thought. Columbia University Press, New York.

Böhm, K. and Durst, S. (2025) "Knowledge Management in the Age of Generative Artificial Intelligence – From SECI to GRAI", VINE Journal of Information and Knowledge Management Systems, Early Cite, pp. 1-16.

Bratianu, C. (2025) "Artificial Knowledge and Its Challenges for the Knowledge Management Systems", Business Excellence and Management, Vol. 15, Special Issue 5, pp. 257-268.

Bratianu, C. and Bejinaru, R. (2019) "The Theory of Knowledge Fields: A Thermodynamics Approach", Systems, Vol. 7, No. 2, pp. 1-12.

Bratianu, C. and Bejinaru, R. (2020) "Knowledge Dynamics: A Thermodynamics Approach", Kybernetes, Vol. 49, No. 1, pp. 6-21.

Davenport, T.H. and Prusak, L. (2000). Working Knowledge: How Organizations Manage what they Know. Harvard Business School Press, Boston, MA.

Di Vaio, A., Latif, B., Gunarathne, N., Gupta, M. and D'Adamo, I. (2024) "Digitalization and Artificial Knowledge for Accountability in SCM: A Systematic Literature Review", Journal of Enterprise Information Management, Vol. 37, No. 2, pp. 606-672.

Gregor, S. and Hevner, A. (2013) "Positioning and Presenting Design Science Research for Maximum Impact", MIS Quarterly, Vol. 37, pp. 337-355.

Hadjimichael, D. and Tsoukas, H. (2019) "Towards a Better Understanding of Tacit Knowledge in Organizations: Taking Stock and Moving Forward", Academy of Management Annals, Vol. 13, pp. 672-703.

Harfouche, A., Quinio, B., Saba, M. and Saba, P.B. (2022) "The Recursive Theory of Knowledge Augmentation: Integrating Human Intuition and Knowledge in Artificial Intelligence to Augment Organizational Knowledge", Information Systems Frontiers, Vol. 26, October 2022, pp. 1-16.

Hevner, A.R., March, S.T., Park, J. and Ram, S. (2004) "Design Science in Information System Research", MIS Quarterly, Vol. 28, pp. 75-105.

Jarrahi, M.H., Askay, D., Eshraghi, A. and Smith, P. (2023) "Artificial Intelligence and Knowledge Management: A Partnership Between Human and AI. Business Horizons, Vol. 66, No. 1, pp. 87-99.

Kolb, D.A. (2015). Experiential Learning: Experience as the Source of Learning and Development. Second edition. Pearson, Upper Saddle River.

Konno, N. (2024). Kōsō-ryoku: Conceptualizing Capability. For Innovation and Management in the Age of Para-existence. Springer Nature Singapore, Singapore.

Kurzweil, R. (2024). The Singularity Is Near: When We Merge with AI. The Bodley Head, London.

Lakoff, G. and Johnson, M. (1999). Philosophy in the Flash: The Embodied Mind and Its Challenges to Western Thought. Basic Books, New York.

Lee, K.F. and Qiufan, C. (2024). Ten Versions for Our Future. Penguin Random House, London.

Nonaka, I. (1994) "A Dynamic Theory of Organizational Knowledge Creation", Organization Science, Vol. 5, No. 1, pp. 14-37.

Nonaka, I. and Takeuchi, H. (1995). The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation. Oxford University Press, Oxford.

Nonaka, I. and Takeuchi, H. (2019). The Wise Company: How Companies Create Continuous Innovation. Oxford University Press, Oxford.

Nonaka, I., Toyama, R. and Hirata, T. (2008). Managing Flow: A Process Theory of Knowledge-Based Firm, Palgrave Macmillan, London.

Pinker, S. (2008). The Stuff of Thought: Language as a Window into Human Nature. Penguin Books, New York.

Polanyi, M. (1983). The Tacit Dimension. Peter Smith, Gloucester.

Russell, S and Norvig, P. (2023). Artificial Intelligence: A Modern Approach. Fourth edition. Pearson, Upper Saddle River.

Simon, H. (1996). The Sciences of the Artificial. Third edition. The MIT Press. Boston.

Tsoukas, H. (2011) "How Should We Understand Tacit Knowledge? A Phenomenological View". In: Easterby-Smith, M. and Lyles, M.A. (Eds.). Handbook of Organizational Learning and Knowledge Management (pp. 453-476). Second Edition. John Wiley & Sons, New York.

Generative AI in Fashion: How Marketing Experts Integrate GenAI in their Knowledge Intensive Work

Nima Taraghi[1], Kathrin Kirchner[2], Ettore Bolisani[1], Tomas Cherkos Kassaneh[1], Enrico Scarso[1]

[1]Department of Management and Engineering, University of Padua, Vicenza, Italy

[2]Department of Engineering Technology, Technical University of Denmark, Ballerup, Denmark

Abstract

The rapid diffusion of Generative Artificial Intelligence (GenAI) is deeply transforming knowledge-intensive work, thanks to its ability not only to automate repetitive tasks but also, and more importantly, to enhance workers' capabilities for creating new knowledge. This is particularly evident in creative sectors where novelty is critical and new knowledge is continuously generated to provide fresh stimuli to the market daily, as in the fashion industry. Despite this, limited empirical research has examined how marketing professionals integrate GenAI into their work. This study addresses this gap by investigating the factors that affect the adoption and use of GenAI among fashion marketing professionals, who are tasked daily with generating innovative ideas to engage potential customers and guide them towards purchasing their company's products. To this aim, the study applies the Diffusion of Innovation Theory as a conceptual interpretative lens to examine the integration of GenAI among 30 purposely selected fashion marketing professionals. This theory, one of the frameworks that current research is using to investigate GenAI adoption processes, assumes that the diffusion of a new technology is influenced by five factors: relative advantage, compatibility, complexity, trialability, and observability. Data were gathered through semi-structured interviews and thematically analysed using MAXQDA software. The analytical process used a hybrid coding strategy, in which the five theory factors provided the deductive structure, while inductive themes were developed to capture the specific knowledge implications of human-AI collaboration. The findings of the investigation shed light on the current usage, benefits, and challenges of employing GenAI to support the knowledge-intensive work of fashion marketing experts. From an academic perspective, it confirms the usefulness of the Diffusion of Innovation framework for investigating the adoption of this new technology. At the same time, it offers valuable insights for marketing experts on the use of GenAI to support their professional activities. The usual limitations of interview-based studies regarding the generalizability of results are acknowledged.

Keywords: Generative Artificial Intelligence, New Technology Adoption, Innovation Framework, Knowledge Workers, Fashion Marketing.

Paper type: Academic Research Paper

1 Introduction

The emergence of Generative Artificial Intelligence (GenAI) is rapidly and deeply changing the landscape of knowledge work, thus affecting the long-term competitiveness of companies (Bharti et al., 2025). Both practitioners and scholars widely acknowledge that GenAI is influencing cognitive work (Noy and Zhang, 2023; Harreis et al., 2023), not only by automating repetitive tasks, but also, and more importantly, by enhancing workers' capabilities, particularly in the creation of new knowledge (Raisch and Krakowski, 2021).). This is why the literature increasingly uses the term "human-AI collaboration" to describe the active and participative nature of GenAI's contribution to knowledge creation (Brynjolfsson et al., 2025). The resulting knowledge can be viewed as a dynamic capability that, by integrating the capacities of both humans and generative systems, enables firms to survive and thrive in increasingly complex environments. Consequently, latecomers risk losing their competitive edge in such an evolving landscape. This is particularly true in creative sectors, where new knowledge is continuously generated to provide fresh stimuli to the market, as in the fashion industry. Evidence of this can be seen in the rapid and widespread adoption of GenAI tools by fashion companies (Amed et al., 2024). In view of the above, the present study aims to investigate the factors that affect the adoption and use of GenAI as an innovative technology in the fashion industry. Our focus is on professionals in marketing, as their work is knowledge-intensive. In fact, they are tasked with generating innovative ideas daily that engage potential customers and guide them to purchase the products of the company they represent.

GenAI supports these professionals by streamlining key cognitive tasks, including automated content creation, concept visualization, product design, and hyper-personalization (Harreis et al., 2023; Park and Ahn, 2024; Taraghi et al., 2025). This research employs a qualitative methodology grounded in Rogers' (2003)

Diffusion of Innovation model to examine the integration of GenAI among 30 purposively selected fashion marketing professionals. In particular, the study addresses the following research question:

RQ: How do fashion marketing experts perceive and adopt Generative AI in their knowledge-intensive work practices?

2 Background of the study

2.1 Generative AI and Knowledge Work

Generative AI is reshaping knowledge work by extending automation beyond routine tasks to complex cognitive activities. Unlike traditional AI systems, which primarily support structured decision-making, GenAI enables the creation of new content, ideas, and solutions, thus directly contributing to knowledge generation processes (Raisch and Krakowski, 2021). Recent studies highlight its ability to enhance productivity and quality in knowledge-intensive tasks, particularly in domains such as writing, design, and problem-solving (Noy and Zhang, 2023; Kirchner et al., 2026).

Rather than substituting human expertise, GenAI fosters new forms of human–AI collaboration, in which individuals and intelligent systems jointly participate in creative and analytical processes (Brynjolfsson et al., 2025). In this context, AI acts as a co-creator, augmenting human capabilities and enabling the recombination of knowledge in novel ways. This interaction can be conceptualized as a dynamic capability, allowing organizations to continuously adapt and innovate in increasingly complex environments. GenAI is not only a tool for efficiency, but also a driver of transformation in how knowledge work is performed, particularly in creative and innovation-driven sectors.

2.2 Generative AI in Creative Industries and Fashion

GenAI is gaining rapid traction in creative industries, where the production of novel and engaging content is central to value creation. By enabling the generation of text, images, and design concepts, GenAI supports creative processes such as ideation, storytelling, and visual development, thereby augmenting rather than replacing human creativity (Harreis et al., 2023; Raisch and Krakowski, 2021).

The fashion industry represents a particularly relevant context for the adoption of GenAI. Characterized by fast-paced innovation cycles and high demand for originality, fashion relies heavily on continuous knowledge creation, especially within marketing and design functions. Recent industry reports highlight the growing use of GenAI tools for applications such as automated content creation, concept visualization, trend forecasting, and personalized customer engagement (Amed et al., 2024; Park and Ahn, 2024).

Despite these developments, the integration of GenAI into creative workflows remains uneven, with organizations facing challenges in meeting skill requirements, integrating tools, and aligning with existing practices (De Cremer et al., 2023). However, despite the growing relevance of GenAI in creative industries, there is still a limited understanding of how these technologies are adopted and integrated at the individual level, particularly within knowledge-intensive marketing roles. In particular, the mechanisms through which professionals perceive, evaluate, and incorporate GenAI into their daily practices remain underexplored. This gap highlights the need for a more fine-grained understanding of individual adoption processes, which can be effectively examined through established technology adoption frameworks.

2.3 Technology Adoption and the Diffusion of Innovation Framework

The adoption of emerging technologies has been examined through several frameworks, including Diffusion of Innovation (DOI), Unified Theory of Acceptance and Use of Technology (UTAUT), and Technology-Organization-Environment (TOE) (Huy et al., 2024; Qu and Kim, 2025; Pokar, 2025). Among these, DOI (Rogers, 2003) is particularly suited to this study, as it focuses on how individuals perceive and adopt new technologies. The theory identifies five key attributes influencing adoption: relative advantage, compatibility, complexity, trialability, and observability (Table 1). Building on this framework, the present study applies the five DOI attributes to examine how fashion marketing professionals perceive and adopt GenAI in their work practices.

Table 1: DOI attributes (Rogers, 2003; García-Avilés, 2020; Noy and Zhang, 2023; Harreis et al., 2023)

DOI attribute	Description
Relative advantage	The degree to which GenAI is perceived as offering benefits over existing practices, such as improved efficiency, productivity, and creativity.
Compatibility	The degree to which GenAI aligns with existing values, work practices, and technological infrastructures.
Complexity	The perceived difficulty of understanding and using GenAI, including the required skills and effort.
Trialability	The extent to which GenAI can be tested before full adoption.
Observability	The visibility of the outcomes and results of GenAI use to others, which may be perceived positively or negatively

3 Research Method

This study adopts a qualitative approach, using the DOI framework as its conceptual lens (Rogers, 2003). The five DOI attributes guide the analysis of how individuals, fashion marketing experts, perceive and integrate GenAI into their knowledge-intensive work practices.

Data were collected through semi-structured interviews with 30 fashion industry professionals from 13 countries (Table 2), recruited via purposive sampling on LinkedIn (Etikan et al., 2016). The sample included a range of roles, from traditional marketing positions to emerging AI-related profiles, ensuring diversity in experience and perspectives. Interviews were conducted online via Zoom between March and October 2025, recorded with consent, and anonymized.

The data were analyzed using a hybrid deductive–inductive thematic approach supported by MAXQDA (Fereday and Muir-Cochrane, 2006; Silver and Lewins, 2014). The five attributes provided the initial coding structure, while inductive themes were developed to capture the knowledge implications of GenAI use. The analysis followed the six phases of thematic analysis (Braun and Clarke, 2006), ensuring rigor and consistency (Yin, 2011). Coding was conducted iteratively, with continuous refinement of themes, and representative quotes were used to support each category and ensure analytical rigor.

4 Findings and Analysis

4.1 Respondent's profile

The respondents represent a diverse group of professionals working across different areas of the fashion industry. Their ages range from early-career individuals to highly experienced professionals with several decades in the industry. Many participants are also actively engaged at the intersection of fashion and technology, with experience in digital fashion, AI-driven initiatives, or technology-oriented roles. Regarding tools, respondents reported an extensive use of GenAI applications in their daily work. Text-based tools such as ChatGPT are widely used for content creation, idea generation, and as a general support for organizing information, conducting research, and refining concepts. Visual tools, particularly Midjourney, play a central role in generating images, supporting creative exploration, and developing marketing visuals. More specialized applications were also mentioned, including AI-powered chatbots for customer interaction, sentiment analysis tools for monitoring feedback, and AI-integrated design software such as Clo3D. Some respondents reported using advanced or custom-built solutions tailored to specific business needs, such as AI systems for content automation, sales support, or product design.

The following sections examine how these professionals perceive and use GenAI, structured according to the five attributes of the DOI framework.

Table 2: Profile of respondents

Interview	Gender	Age	Years of experience	Title	Country
Int1	Woman	38	10	Fashion marketing expert	Italy
Int2	Woman	43	20	Business owner, fashion designer	Trinidad and Tobago
Int3	Woman	40	10	Business owner, handicraft designer	Canada
Int4	Woman	43	4.5	Lecturer in Fashion marketing and branding	UK
Int5	Woman	33	8	Conscious denim designer	US
Int6	Woman	56	12	Digital (AI) fashion creator	Canada
Int7	Man	45	5	Digital project manager	Italy

Interview	Gender	Age	Years of experience	Title	Country
Int8	Woman	37	13	Chief operating officer	Italy
Int9	Woman	37	15	Fashion innovation expert	Italy
Int10	Woman	34	14	Business owner, fashion designer	Bulgaria
Int11	Woman	40	4	Fashion journalist	Iran
Int12	Woman	50	25	Visual designer & AI creative director	Switzerland
Int13	Woman	48	20	Digital marketing expert	Austria
Int14	Woman	37	14	Business owner, fashion designer	Egypt
Int15	Man	43	20	Business owner, sports manufacturer	Belgium
Int16	Woman	57	35	Digital fashion designer	US
Int17	Woman	55	38	Business owner, creative designer	Italy
Int18	Woman	54	35	Business owner, creative designer	Italy
Int19	Man	64	40	Business owner, visual creator	Italy
Int20	Woman	33	10	Sale engineer	Italy
Int21	Woman	55	34	Fashion tech founder	US
Int22	Woman	43	20	Fashion designer, GenAI creator	Italy
Int23	Woman	24	1	Fashion marketing expert	Italy
Int24	Woman	37	15	Fashion marketing expert	Canada
Int25	Woman	58	30	3D product developer/ consultant	US
Int26	Woman	27	5	Fashion innovation consultant and designer	Switzerland
Int27	Woman	43	20	Business owner, designer	Portugal
Int28	Woman	25	4	Fashion sales and marketing expert	UK
Int29	Woman	31	10	Fashion e-commerce expert	India
Int30	Woman	43	20	Product design lead, GenAI expert	Portugal

4.2 Relative advantage

Table 3 outlines sub-themes of relative advantage, showing how GenAI improves performance over traditional methods. Frequency reflects how many respondents mentioned each sub-theme, indicating relevance.

Most respondents emphasize that GenAI is primarily valued for boosting efficiency and saving time, thereby making workflows faster and more productive. They also note increased creativity and better customer personalization. Cost and resource optimization are deemed important, too, especially through reduced reliance on traditional processes. Some respondents highlight benefits for business sustainability. Overall, GenAI is perceived to offer advantages across efficiency, creativity, and cost savings.

Table 3: Sub-themes for Relative Advantage

Sub theme	Description	Quotes	Frequency
Efficiency and Time Compression	GenAI accelerates workflows and can enable partial task automation, reducing time and human effort.	"… with AI, your efficiency is improved, your productivity becomes 2x" (Int29) "… what AI can do in a couple of minutes would take humans days or weeks" (Int15)	27
Augmented Creativity and Ideation	GenAI enhances creative processes by supporting ideation, inspiration, and the generation of novel concepts.	"… it's the fastest way of ideation and having, doing a more accurate brainstorm also, getting immediate response or ideas to present to clients" (Int27) "... it's opening up a door to next level creativity" (Int13)	23
Customer Experience and Personalizatio n	GenAI enables more personalized, interactive customer experiences, thereby improving engagement and relevance.	"… AI tool helps us to make more personalized connections with the customer … Make messages more friendly. More personalized for every customer" (Int14) "… analyze what your customer is looking for, and generate descriptions like in real time for the product" (Int9)	22

Sub theme	Description	Quotes	Frequency
Cost and Resource Optimization	GenAI reduces reliance on traditional production resources (e.g., photoshoots, teams), leading to significant cost savings.	"So basically, photo shooting is being replaced and significantly optimized … the savings are substantial in this regard" (Int30) "A third factor is certainly cost, because with the same budget as before, we might now be able to carry out more or more complete activities" (Int18)	20
Sustainability and Process Optimization	GenAI supports more sustainable practices by reducing physical production, waste, and overproduction.	"You can try a lot of things without having to spend money on real products. So, you can use this for tests, and in terms of environmental benefits, it's huge" (Int27)	8

4.3 Compatibility

Table 4 outlines the sub-themes related to compatibility and shows that compatibility is mainly driven by how easily GenAI can be integrated into existing workflows and used as an augmenting tool rather than a full replacement. Respondents emphasize its role as a co-creator, supporting creative and marketing activities while aligning with brand identity and industry-specific requirements. At the same time, compatibility is shaped by organizational readiness, including skills and openness to change, as well as constrained by ethical, legal, and trust-related concerns. The respondents' overall opinion is that compatibility can be missing, as it depends on how well the GenAI use fits with specific work practices, brand values, and the capabilities of both individuals and organizations. Furthermore, ethical, legal, and trust issues afflicting current tools contribute to lowering their compatibility.

Table 4: Sub-themes for Compatibility

Sub theme	Description	Quotes	Frequency
Task, Brand, and Industry Fit	Compatibility depends on how well GenAI aligns with specific work tasks, brand identity, and industry context. This includes its ability to support creative and marketing activities while maintaining brand DNA and adapting to sector-specific expectations.	"… to create images that express a concept an idea that transfer the values of the brand which translates into images" (Int23) "… AI cannot be useful because when you combine the traditional handicraft with modern handicraft, it has specific customers and the niche market is not very big" (Int3)	30
Integration and Augmentation of Existing Practices	GenAI is perceived as compatible when it can be integrated into existing workflows and used as a tool that augments, rather than replaces, human practices. This includes its role as a co-creator that supports and enhances existing expertise.	"AI is just another tool… like Photoshop" (Int24) "This shift, this co-creation and coexistence with the machine can be carried throughout the entire pipeline" (Int18)	29

Table 4: Sub-themes for Compatibility (cont.)

Sub theme	Description	Quotes	Frequency
Organizational Readiness and Capability	Compatibility is influenced by individuals' and teams' readiness to adopt GenAI, including their openness to change, skills, and ability to integrate new technologies into existing practices.	"I think mostly the mindset, because technically, they are easy to learn, it's just that if people don't want to learn them, then, well, they won't, you know" (Int26) "It's a new technology. Not all team can understand it. So I give them more courses to be with me in this transformation …" (Int14)	23
Ethical, Legal, and Trust Barriers	Compatibility is constrained by concerns related to ethics, data privacy, copyright, and customer trust.	"… the privacy, copyright issues are the biggest challenges" (Int8) "... the risk is that you create an image that can be similar to a real existing person and there can be some legal problems …" (Int23)	16

4.4 Complexity

Table 5 illustrates the key points of complexity associated with GenAI use, arising from both the effort required to use it and the uncertainty of its outcomes. Respondents highlight the need for specific skills and continuous learning, as well as difficulties in controlling outputs and achieving reliable, high-quality results. In addition, fragmented workflows and the need to combine multiple tools increase operational complexity. This is further reinforced by a gap between expectations and actual performance, along with ethical and risk-related concerns such as privacy and data security. Overall, complexity reflects both technical and organizational challenges that can limit effective adoption.

Table 5: Sub-themes for Complexity

Sub theme	Description	Quotes	Frequency
Output Uncertainty and Control Limitations	GenAI use is characterized by uncertainty in outputs, limited control over results, and inconsistencies in quality and accuracy.	"… sometimes it's the AI, it's not the right answer for everything. You have to understand where and when and how to use it" (Int27) "… because we all know AI makes so many mistakes right now" (Int21)	20
Skill Intensity and Learning Curve	Although GenAI tools are accessible, effective use requires specific skills (e.g., prompting, domain knowledge) and continuous learning due to rapid technological evolution.	"You need to know how to use it before actually like doing it" (Int28) "So I have some mistakes. Take a time to learn it and integrate it in our projects" (Int14) "… it's not like you are coming from day one and start getting results, nice results, usable results. No, you have to learn. Someone has to teach you" (Int30)	20
Expectation–Reality Gap	There is a mismatch between expectations of AI capabilities and actual performance, leading to frustration and dissatisfaction.	" … one of the things that is very mystifying about AI is that everything is possible. (Int27) "… GNI isn't always capable of replicating a high-quality product. It's still not always possible to create something as good as a filmed video, an image, or a photo" (Int12)	17
Workflow Fragmentation and Tool Ecosystem Complexity	The need to combine multiple tools and manually adjust outputs creates fragmented workflows, reducing overall efficiency.	"… for someone like me, with a limited budget, one of the biggest difficulties is gaining access to truly professional tools, which often come with very high costs" (Int22) "Another major challenge … is the overwhelming number of tools available, and the ability to choose which ones to use" (Int18)	13
Ethical and Risk-Related Complexity	Concerns related to privacy, intellectual property, and misuse of AI outputs add an additional layer of complexity to adoption.	"The biggest challenge definitely has two sides: one is ethical, and the other is legal—it's the legal side" (Int18) "… the challenge… is the privacy and data security, because we know very well what you share with ChatGPT" (Int8)	11

4.5 Trialability

Table 6 summarizes trialability sub-themes, showing how GenAI can be tested before full adoption. GenAI denotes strong trialability due to its ease and speed, enabling quick idea testing, prototyping, and low-cost experimentation. Respondents emphasize flexible, iterative testing as a key benefit. However, scaling to professional use introduces challenges, such as maintaining brand consistency, quality, and control. Reliability and output limitations can also reduce effectiveness. According to the experts, trialability is high early on but more limited in large-scale applications.

Table 6: Sub-themes for Trialability

Sub theme	Description	Quotes	Frequency
Ease and Accessibility of Experimentation	GenAI enables fast, low-cost, and low-effort experimentation, allowing users to rapidly prototype ideas, explore multiple alternatives, and test concepts with audiences. This encourages iterative exploration in the early stages.	"… the key benefit of AI in marketing, it's a fast testing of ideas, flexibility of the process" (Int24) "Let's say we like to consider ourselves pioneers—not because AI itself is new (we know well that it began decades ago), but because its wide adoption has really accelerated only in recent years" (Int17)	20
Iterative Testing and Creative Exploration	Users can continuously refine, adjust, and iterate on output, enabling dynamic experimentation and creative exploration across multiple variations and ideas.	"… the key benefits, definitely the flexibility of the process. You can adjust, redo, rethink in seconds " (Int24) "And ever since then, I've just been like, really, you know, playing around with it, trying with … all the new tools that are coming out" (Int16)	18
Reliability and Validity Limitations of Experimentation	The outcomes of GenAI experimentation are often limited by issues of quality, realism, and consistency, reducing the reliability of testing results and making it difficult to draw actionable conclusions.	"There are some try-on tools that are really good on some kind of clothes but really bad on some others … the technology is not enough developed" (Int23) "… even though you're getting multiple attempts, I've found it still is not as accurate as I would like it to be. So in the end, I unsubscribe from that software" (Int2)	16
Constraints in Scaling Experimentation	While initial experimentation is easy, applying GenAI in professional and large-scale contexts becomes more complex due to brand constraints, organizational requirements, and the difficulty of maintaining consistency and control.	"Post-production is still necessary because even the best models, despite being guided and trained over the years … still aren't able to deliver exactly what's needed" (Int12) "… but then it began to work very good and in the late two years things have changed constantly day by day and the results are pretty different" (Int19)	15

4.6 Observability

Table 7 presents the sub-themes associated with observability, focusing on how the outcomes of GenAI use become visible to others. External observability of results obtained with GenAI is not straightforward because it depends on the visibility of GenAI use and its outcomes to others, which can attract significant attention and curiosity. Respondents highlight that AI-related content often gains visibility due to its novelty, while increasing realism and detectability shape whether users recognize AI involvement. Once visible, these outcomes are interpreted differently by audiences, leading to both positive and negative reactions depending on perceptions of authenticity and expectations. In sum, observability in this context encompasses not only the visibility of results but also how others perceive and evaluate them.

Table 7: Sub-themes for Observability

Sub theme	Description	Quotes	Frequency
Visibility and Attention Effects	The use of GenAI increases visibility and attracts attention, as the novelty and "AI label" generate curiosity, media coverage, and audience engagement.	"I've also used a few posts created by AI on my LinkedIn, which had over 80,000 views" (Int15) "…. when a person or a company will start to use it and then they will see the benefits and the benefits will be shared to like the big public let's say then after everyone will start to use it" (Int28)	20

Sub theme	Description	Quotes	Frequency
Detectability and Transparency of AI Use	Observability depends on the extent to which AI use is recognizable, either through the detectability of AI-generated outputs or through explicit disclosure.	"I mean, the beginning, it was very difficult to have a video that looked natural. Now people don't even know the difference between an AI, I can't even tell myself" (Int16) "They have gone full gen AI on everything and it does look slightly strange and weird and its kind of you get that uncanny valley feel from it" (Int4)	18
Interpretation of Visible Outcomes	Once visible, GenAI outputs are interpreted differently by audiences, leading to varied reactions depending on perceived realism, authenticity, and expectations.	"Customers didn't take it quite well because they felt somehow betrayed by the brand choosing to avoid humans as models" (Int7) "The response was very positive and the client itself felt like it was also working on that because we were tweaking the outfits" (Int27)	18

5 Discussion and Conclusion

Fashion marketing professionals perceive GenAI as a technology that significantly enhances efficiency, especially content generation and ideation (Noy and Zhang, 2023). The high frequency of efficiency- and creativity-related benefits indicates that they primarily adopt GenAI to accelerate workflows and expand creative output, but the value is realized only when the technology can be effectively embedded within existing workflows and aligned with brand identity and creative standards where GenAI is an active co-creator partner and enable a combination of human creativity and machine efficiency (Raisch and Krakowski, 2021). This explains why compatibility-related factors, particularly integration and task–brand fit, are among the most important in respondents' opinions.

The findings also reveal a fundamental tension between ease of access and depth of use. While GenAI is highly accessible and supports rapid experimentation, effective and sustained use requires developing new forms of knowledge, and experts must learn to guide outputs, evaluate quality, and manage iterative processes. This indicates that knowledge work is shifting from creation to orchestration (Raisch and Krakowski, 2021) and . explains the strong presence of complexity-related sub-themes, particularly those linked to skill requirements and output uncertainty. Therefore, adoption is not simply about using a new tool, but about developing the capabilities needed to control and integrate it into professional practice (Kirchner et al., 2025). We can also see this tension in trialability, where experimentation is easy and accessible, while challenges related to quality arise when scaling up use. As a result, adoption is not a linear progression from trial to full use, but a staged process in which practical constraints moderate early enthusiasm. In the context of digital fashion marketing, GenAI outputs are highly observable, which increases awareness and attention. However, unlike traditional assumptions in diffusion theory, this visibility does not consistently reinforce adoption. Instead, interpretations of visible outputs vary, with both positive and negative reactions depending on perceived authenticity and quality. This suggests that adoption is influenced not only by functional performance but also by how outcomes are socially interpreted and evaluated (Brüns and Meißner, 2024).

As a conclusion, the adoption of GenAI in fashion marketing is best understood as a dynamic process. Professionals adopt GenAI to enhance productivity and creativity. Still, its effective use depends on its integration into existing knowledge systems, the development of new skills, and the management of audience perceptions. This highlights that GenAI does not simply improve knowledge work but reshapes it, requiring a shift in how knowledge is created, controlled, and evaluated.

6 Limitations, Implications, and Future Directions

From a theoretical perspective, the findings extend the DOI framework by showing that attributes such as observability and trialability are more dynamic and context-dependent in the case of GenAI.

From a practical perspective, organizations should view GenAI not only as a productivity tool but also as a capability that requires new skills and organizational adaptation. They should invest in developing employees' AI-related skills, particularly in guiding and evaluating outputs. GenAI should be strategically integrated into existing workflows through a phased approach, starting with pilot use cases. Strong quality control and governance mechanisms, supported by human oversight, are essential to ensure accuracy and alignment with brand identity. While experimentation is relatively easy, scaling requires addressing reliability and integration challenges, suggesting a staged adoption from experimentation to controlled deployment and broader

integration. Finally, firms must carefully manage customer perception and trust by adopting clear communication strategies and ensuring appropriate transparency in customer-facing applications.

This study has limitations, including its industry-specific focus and participants' early-stage GenAI adoption, which may limit generalizability. Future research could extend these insights through cross-industry and longitudinal studies.

7 Ethics declaration

Participation was voluntary, with informed consent obtained from all participants. No sensitive data was collected, and all responses were anonymized and handled in accordance with data protection regulations. Formal ethics approval was not required under institutional guidelines.

8 AI declaration

We used ChatGPT 5.3 to revise and correct the written text and rename the theme's label.

References

Amed, I., Berg, A., Balchandani, A., Rölkens, F., and Barrelet, D. (2024) *The state of fashion 2024*, [online], McKinsey & Company, https://www.mckinsey.com/industries/retail/our-insights/state-of-fashion-2024.

Bharti, I., Chauhan, K. and Aggarwal, P. (2025) Generative AI: Next frontier for competitive advantage". In: Bharti, I., Chauhan, K. and Aggarwal, P. (eds), *Enhancing communication and decision-making with A*I, IGI Global, New York, pp 1-36.

Braun, V. and Clarke, V. (2006) Using thematic analysis in psychology, *Qualitative research in psychology*, Vol. 3, No. 2, pp 77-101.

Brüns, J.D. and Meißner, M. (2024) Do you create your content yourself? Using generative artificial intelligence for social media content creation diminishes perceived brand authenticity, *Journal of Retailing and Consumer Services*, Vol. 79, p 103790.

Brynjolfsson, E., Li, D. and Raymond, L. (2025) Generative AI at work, *The Quarterly Journal of Economics*, Vol. 140, No, 2, pp 889-942.

De Cremer, D., Bianzino, N.M. and Falk, B. (2023) How generative AI could disrupt creative work, *Harvard Business Review*, Vol, 13, p 13.

Etikan, I., Musa, S.A. and Alkassim, R.S. (2016) Comparison of convenience sampling and purposive sampling, *American journal of theoretical and applied statistics*, Vol. 5, No.1, pp 1-4.

Fereday, J. and Muir-Cochrane, E. (2006) Demonstrating rigor using thematic analysis: A hybrid approach of inductive and deductive coding and theme development, *International journal of qualitative methods*, Vol. 5, No. 1, pp 80-92.

García-Avilés, J.A. (2020) Diffusion of innovation. In: van den Bulck, J., Sharrer, E., Ewoldsen, D. and Mares, M-L. (Eds) *The International Encyclopedia of media psychology*, Vol. 1.

Harreis, H., Koullias, T., Roberts, R., and Te, K. (2023) "Generative AI: Unlocking the future of fashion", [online], McKinsey & Company, https://www.mckinsey.com/industries/retail/our-insights/generative-ai-unlocking-thefuture-of-fashion/.

Huy, L. V., Nguyen, H. T., Vo-Thanh, T., Thinh, N. H. T., and Thi Thu Dung, T. (2024) Generative AI, why, how, and outcomes: A user adoption study, *AIS Transactions on Human-Computer Interaction*, Vo. 16, No. 1, pp 1-27.

Kirchner, K., Bolisani, E., Kassaneh, T.C., Scarso, E. and Taraghi, N. (2025) Impact of Work Experience on the Use of Generative AI for Knowledge Management, *20th International Forum on Knowledge Asset Dynamics*, pp 1518-1525.

Kirchner, K., Bolisani, E., Kassaneh, T.C., Scarso, E. and Taraghi, N. (2026) Bug or feature? Investigating the impact of generative AI on knowledge creation in software engineering, *VINE Journal of Information and Knowledge Management Systems*, ahead-of-print

Noy, S., and Zhang, W. (2023) Experimental evidence on the productivity effects of generative artificial intelligence, *Science*, Vol. 381, No. 6654, pp 187-192.

Park, J., and Ahn, S. (2024) Traditional vs. AI-generated brand personalities: Impact on brand preference and purchase intention, *Journal of Retailing and Consumer Services*, Vol. 81, 104009.

Pokar, J. (2025) Unlocking the potential of generative artificial intelligence: adoption drivers and organisational performance, *International Journal of Business Innovation and Research*, Vol. 38, No. 9, pp 1-23

Qu, C. and Kim, E. (2025) Investigating AI adoption, knowledge absorptive capacity, and open innovation in Chinese apparel MSMEs: An extended TAM-TOE model with PLS-SEM analysis, *Sustainability*, Vol. 17, No. 5, p 1873.

Rogers, E. M. (2003) *Diffusion of innovations* (5th ed.). New York, NY: Free Press

Raisch, S., and Krakowski, S. (2021) Artificial intelligence and management: The automation–augmentation paradox, *Academy of Management Review*, Vol. 46, No. 1, pp 192-210.

Silver, C., and Lewins, A. (2014) *Using software in qualitative research: A step-by-step guide*, Sage Publications, Thousand Oaks.

Taraghi, N., Scarso, E., and Kirchner, K. (2025) Reshaping Customer Experience with Generative AI: Insights from Fashion Experts, *Proceedings EMAC Fall Conference 2025*.

Yin, R.K. (2011), *Applications of case study research*, Sage publications, Newbury Park, CA.

The Hidden Cost of Flexibility: Knowledge Visibility in Flexible Work Arrangements

Elina Pentjusa, Elina Gaile-Sarkane
Riga Technical University, Riga, Latvia

Abstract

Flexible work arrangements have become a defining feature of contemporary organizations, yet their implications for organizational knowledge processes remain insufficiently theorized. Existing research has primarily examined flexible work through employee autonomy, wellbeing, productivity, retention, and work-life balance. While these perspectives are important, they provide only a partial account of how flexible work reshapes the creation, sharing, retention, and renewal of organizational knowledge. This paper argues that flexible work should be understood not only as a human resource management or work design issue, but also as a knowledge governance issue.

The paper develops a conceptual framework explaining how flexible work arrangements influence knowledge visibility in organizations. Flexible work is treated as a multidimensional construct comprising spatial, temporal, and organizational flexibility rather than as a single uniform practice. Drawing on tacit knowledge theory, organizational knowledge creation, situated learning, communities of practice, technology affordance theory, transactive memory systems, and knowledge governance, the paper argues that the knowledge effects of flexibility are conditional. Flexible work may strengthen knowledge processes by supporting autonomy, focused work, digital traceability, and access to distributed expertise, but may also weaken tacit knowledge transfer, newcomer learning, shared interpretation, and expertise visibility when knowledge processes are not intentionally governed.

The main contribution is the development of a flexibility-knowledge visibility framework. The framework explains how different forms of flexibility affect organizational knowledge through four mechanisms: tacit knowledge transfer, digital and social visibility, transactive memory, and participation in learning communities. The paper develops five propositions for future empirical research and outlines governance mechanisms for maintaining knowledge visibility in flexible work contexts. The argument is not that flexibility is inherently harmful. Rather, knowledge fragmentation emerges when organizations redesign where and when work occurs without redesigning how knowledge becomes visible, accessible, interpretable, and renewable.

Keywords: Flexible work arrangements; knowledge visibility; tacit knowledge; hybrid work; knowledge governance

Paper type: Academic Research Paper

1 Introduction

Flexible work arrangements have become a lasting feature of contemporary organizational life. Remote work, hybrid work, flexible schedules, and digitally mediated collaboration have changed not only where and when work is performed, but also how employees interact, learn, coordinate expertise, and participate in organizational knowledge processes. Much of the existing discussion has examined flexible work through outcomes such as autonomy, job satisfaction, productivity, work-life balance, and employee retention (Allen, Golden, & Shockley, 2015; Gajendran & Harrison, 2007; Kossek & Lautsch, 2018). These perspectives are important, but they do not fully explain how flexible work affects the visibility, transfer, and renewal of organizational knowledge.

This gap is significant because organizational knowledge is not limited to codified information. It also includes tacit knowledge, situated judgement, shared routines, relational understanding, and knowledge of who holds relevant expertise (Nonaka, 1994; Polanyi, 1966; Wegner, 1987). Such knowledge is often developed through observation, informal correction, peripheral participation, and repeated interaction in communities of practice (Brown & Duguid, 1991; Lave & Wenger, 1991; Wenger, 1998). When work becomes spatially distributed, temporally asynchronous, and digitally mediated, these knowledge processes do not disappear, but their conditions change.

This paper argues that flexible work should be understood as a knowledge governance issue. The central question is not whether flexible work is good or bad for organizations, but under what conditions flexibility supports or weakens knowledge creation, sharing, retention, and renewal. Flexible work may improve autonomy, reduce interruptions, support focused work, and increase access to distributed expertise. At the same time, it may reduce tacit learning, weaken ambient awareness, fragment shared context, and make

expertise less visible if knowledge processes are left to emerge informally (Cheng, Sun, Zhong, & Li, 2023; Hesselbarth, Alfes, & Festing, 2024; Majchrzak, Faraj, Kane, & Azad, 2013).

The paper addresses the following research question: How do flexible work arrangements reshape organizational knowledge visibility, and what governance mechanisms are required to protect knowledge creation, sharing, retention, and renewal in flexible work contexts?

To answer this question, the paper develops a conceptual framework that treats flexible work as a multidimensional construct. Spatial flexibility concerns where work is performed, temporal flexibility concerns when work is performed, and organizational flexibility concerns how work is coordinated, mediated, and governed. This distinction is important because different forms of flexibility have different implications for tacit knowledge transfer, digital visibility, expertise coordination, and newcomer learning (Allen et al., 2015; Hesselbarth et al., 2024).

The paper contributes by shifting flexible work research from individual-level outcomes toward organizational knowledge visibility. It also extends knowledge management literature by explaining how flexible work changes the socio-technical conditions through which knowledge becomes visible, transferable, and governable.

2 Theoretical Background

2.1 Flexible work arrangements as a multidimensional construct

Flexible work arrangements refer to work practices that provide discretion over where, when, and how work is performed. The concept includes remote work, hybrid work, flexible schedules, asynchronous collaboration, compressed working time, and digitally mediated coordination. However, treating flexibility as a single category is analytically weak because different forms of flexibility produce different organizational consequences (Allen et al., 2015; Kossek & Lautsch, 2018).

This paper distinguishes between spatial, temporal, and organizational flexibility. Spatial flexibility concerns where work is performed, temporal flexibility concerns when work is performed, and organizational flexibility concerns how work is coordinated through roles, routines, platforms, and managerial practices. This distinction is central because each dimension creates different knowledge risks: spatial flexibility may reduce informal observation, temporal flexibility may weaken shared rhythms, and organizational flexibility may either reduce or intensify fragmentation depending on the quality of knowledge systems and governance routines.

Recent research supports the need for this more differentiated view. Cheng et al. (2023) show that flexible work arrangements influence knowledge sharing through mechanisms such as workplace loneliness and task interdependence. Hesselbarth et al. (2024) argue that technology-driven work arrangements should be understood from a complexity perspective rather than as simple work location choices. Eng, Aboagye, Bergsten, and Strömberg (2025) further show that flexibility may affect innovation, productivity, and organizational attractiveness differently depending on organizational context.

2.2 Organizational knowledge as tacit, social, and technological

Organizational knowledge includes explicit knowledge that can be documented, tacit knowledge that is embedded in experience, and social knowledge that is sustained through relationships and practice (Nonaka, 1994; Polanyi, 1966; Spender, 1996). In knowledge-intensive organizations, performance depends not only on access to information, but also on interpretation, judgement, shared context, and the ability to locate relevant expertise (Grant, 1996; Nonaka & Takeuchi, 1995).

Knowledge visibility describes the extent to which knowledge, expertise, reasoning, and learning opportunities are accessible and interpretable within an organization. It is broader than information availability: a document may be accessible without being understood, a decision may be recorded without its reasoning being visible, and expertise may exist without being identifiable to those who need it. Thus, knowledge visibility concerns the social and technical conditions that make knowledge locatable, interpretable, and applicable.

Flexible work changes knowledge visibility in two opposing ways. Digital tools can increase visibility by creating searchable records, shared repositories, project histories, and communication traces (Leonardi, 2014; Treem & Leonardi, 2013). At the same time, flexible work may reduce visibility when interaction moves into private channels, when employees have fewer informal encounters, or when digital communication captures information but not context (Majchrzak et al., 2013). The central issue is therefore not the presence or absence of technology, but whether technology and social routines jointly support knowledge interpretation.

2.3 Tacit knowledge, social learning, and flexible work

Tacit knowledge is difficult to formalize because it is rooted in experience, practice, and situated judgement (Polanyi, 1966). Nonaka's (1994) theory of organizational knowledge creation explains that knowledge develops through interaction between tacit and explicit forms. The SECI model identifies socialization, externalization, combination, and internalization as key processes through which organizational knowledge is created and renewed (Nonaka & Takeuchi, 1995).

Situated learning theory adds a further dimension. Lave and Wenger (1991) argue that learning occurs through participation in social practice. Newcomers develop competence through legitimate peripheral participation, gradually moving from observation to fuller involvement. Communities of practice sustain knowledge through shared repertoires, mutual engagement, and collective interpretation (Brown & Duguid, 1991; Wenger, 1998). In flexible work contexts, these learning processes may be weakened if employees, especially newcomers, lose access to informal observation, shadowing, and low-threshold questions.

However, co-presence is not the only basis for learning. Digital communities, virtual mentoring, recorded demonstrations, transparent discussion channels, and structured reflection routines can support learning across distance when they are deliberately connected to real work practices (Leonardi, 2014; Majchrzak et al., 2013; Wenger, 1998).

2.4 Digital affordances and transactive memory

Digital work technologies shape knowledge processes through affordances, understood as possibilities for action that emerge from the relationship between users and technology (Majchrzak et al., 2013; Treem & Leonardi, 2013). Collaboration platforms may support knowledge sharing by making conversations searchable, decisions traceable, and expertise networks more accessible.

Yet digital visibility is selective. It may make documents and messages visible while leaving informal reasoning, judgement, and contextual cues hidden. Although digital tools can support knowledge sharing by making communication visible, their effect depends on how employees use technology and how organizations structure interaction (Leonardi, 2014). In flexible work settings, digital systems may therefore create either knowledge integration or knowledge fragmentation.

Transactive memory systems theory explains why expertise visibility is critical. Teams perform more effectively when members know who knows what and how to access relevant expertise (Lewis, 2003; Ren & Argote, 2011; Wegner, 1987). Flexible work may strengthen transactive memory by making expertise searchable across locations. It may also weaken transactive memory if employees have fewer opportunities to observe who solves problems, who is trusted, and who holds practical experience beyond formal job descriptions.

2.5 Knowledge governance in flexible work

Knowledge governance refers to the organizational mechanisms through which knowledge creation, sharing, transfer, and use are shaped by structures, routines, incentives, technologies, and managerial practices (Foss, Husted, & Michailova, 2010; Grandori, 2001; Heiman & Nickerson, 2004). This perspective is important because knowledge outcomes are not produced by flexible work arrangements alone. They depend on how organizations design the conditions for interaction, documentation, learning, and expertise coordination.

In flexible work contexts, knowledge governance becomes more important because informal coordination can no longer be assumed. Managers must decide which knowledge should be documented, which interactions require synchronous discussion, which learning processes require social exposure, and which tools should be used for different knowledge purposes. Governance therefore refers to deliberate knowledge design rather than excessive control.

This paper therefore defines knowledge governance for flexible work as the intentional design of social and technical mechanisms that make organizational knowledge visible, accessible, interpretable, and renewable across distributed work arrangements.

3 Methodology

This paper is designed as a conceptual academic research paper based on an integrative literature review and theory-building logic. A conceptual approach is appropriate when the aim is to clarify constructs, connect fragmented research streams, and develop propositions for future empirical research (Torraco, 2005; Whetten, 1989). The purpose is to develop a theoretically grounded framework explaining how flexible work changes the conditions under which organizational knowledge becomes visible, transferable, and renewable.

The review process was concept-driven rather than purely systematic. Literature was selected from five interrelated domains: flexible work arrangements, remote and hybrid work, knowledge management, organizational learning, and digital work technologies. Priority was given to peer-reviewed journal articles, recent post-pandemic studies, and foundational theoretical works that remain central to knowledge management and organizational learning. The inclusion logic followed three criteria. First, sources were included if they helped define flexible work as a multidimensional construct. Second, sources were included if they explained knowledge processes such as tacit knowledge transfer, knowledge sharing, situated learning, knowledge creation, and expertise coordination. Third, sources were included if they addressed mechanisms through which technology, distance, flexibility, or coordination affect knowledge visibility.

The analytical process involved four stages. The first stage was construct clarification. Flexible work arrangements were separated into spatial, temporal, and organizational flexibility in order to avoid treating flexibility as a single uniform variable (Allen et al., 2015; Kossek & Lautsch, 2018). The second stage was mechanism identification. The literature was examined to identify knowledge mechanisms that are likely to be affected by flexible work. Four mechanisms were selected as central to the framework: tacit knowledge transfer, digital and social visibility, transactive memory, and participation in learning communities.

The third stage was theoretical integration. Foundational knowledge theories were combined with contemporary digital work literature to explain different aspects of knowledge visibility: tacit and explicit knowledge (Nonaka, 1994; Nonaka & Takeuchi, 1995; Polanyi, 1966), situated learning and communities of practice (Brown & Duguid, 1991; Lave & Wenger, 1991; Wenger, 1998), digital affordances (Majchrzak et al., 2013; Treem & Leonardi, 2013), transactive memory systems (Lewis, 2003; Wegner, 1987), and knowledge governance (Foss et al., 2010; Grandori, 2001).

The fourth stage was framework development. The flexibility-knowledge visibility framework was developed by comparing the potential benefits of flexible work with the risks it creates for knowledge visibility. The resulting propositions are conditional rather than deterministic because the effects of flexibility depend on the form of flexibility, the nature of work, digital affordances, and knowledge governance mechanisms.

4 Conceptual Framework and Propositions

The conceptual framework explains how flexible work arrangements influence organizational knowledge visibility. It starts from the premise that flexible work does not directly determine knowledge outcomes; rather, it changes the socio-technical conditions under which knowledge is created, shared, interpreted, retained, and renewed.

The framework links three dimensions of flexibility - spatial, temporal, and organizational - to four knowledge mechanisms: tacit knowledge transfer, digital and social visibility, transactive memory, and participation in learning communities. These mechanisms shape outcomes such as knowledge integration, fragmentation, newcomer learning, innovation capacity, and organizational memory. Knowledge governance is positioned as the key moderating condition because it determines whether flexible work is supported by routines, technologies, and managerial practices that preserve knowledge visibility (Foss, Husted, & Michailova, 2010; Grandori, 2001; Heiman & Nickerson, 2004). The proposed framework is summarized in Table 1.

Table 1: Flexible work dimensions and knowledge governance responses

Flexible work dimension	Knowledge risk	Core mechanism	Governance mechanism
Spatial flexibility	Loss of tacit learning	Reduced observation	Mentoring, shadowing, purposeful co-presence
Temporal flexibility	Weaker shared interpretation	Delayed feedback	Decision logs, feedback cycles, synchronous discussion
Organizational flexibility	Tool-based fragmentation	Dispersed knowledge systems	Integrated platforms, documentation standards
Flexible work dimension	Knowledge risk	Core mechanism	Governance mechanism
High task interdependence	Expertise opacity	Weaker transactive memory	Expert directories, project histories, communities of practice
Newcomer integration	Formal but shallow learning	Reduced peripheral participation	Hybrid onboarding, guided participation, real work exposure

Table 1 summarizes the central logic of the framework. Spatial flexibility threatens tacit learning, temporal flexibility affects shared interpretation, and organizational flexibility creates risks when digital tools multiply

information without preserving context. High task interdependence may weaken transactive memory, while newcomer integration becomes vulnerable when onboarding is detached from real work participation. These risks can be reduced through governance mechanisms that reconnect flexibility with observation, interpretation, documentation, expertise visibility, and situated learning.

The framework rests on three assumptions. First, flexible work is multidimensional and should not be collapsed into a single variable because spatial, temporal, and organizational flexibility affect different knowledge mechanisms (Allen, Golden, & Shockley, 2015; Hesselbarth, Alfes, & Festing, 2024). Second, knowledge visibility is not the same as information availability: digital tools may increase access to information while leaving meaning, context, and expertise unclear (Leonardi, 2014; Majchrzak, Faraj, Kane, & Azad, 2013). Third, the relationship between flexibility and knowledge outcomes is conditional. Flexible work may support knowledge integration when governance mechanisms are strong, but generate fragmentation when organizations rely on spontaneous knowledge flows under distributed conditions.

4.1 Spatial flexibility and tacit knowledge transfer

Spatial flexibility changes where work is performed and therefore affects access to proximity-based learning. In co-located settings, employees often learn by observing how experienced colleagues interpret problems, handle exceptions, and make situated decisions. This type of learning is central to tacit knowledge transfer because tacit knowledge is practical, experiential, and difficult to fully codify (Nonaka, 1994; Polanyi, 1966). When employees work remotely or in hybrid arrangements, these informal learning opportunities may become less frequent and less visible.

The risk is strongest where work requires judgement, exception handling, or high task interdependence. Since competence develops through participation in practice, not only formal instruction, reduced access to observation, shadowing, and informal correction may weaken practical knowledge transfer (Lave & Wenger, 1991).

Proposition 1: Spatial flexibility is more likely to weaken tacit knowledge transfer when work requires situated judgement, informal observation, and high task interdependence.

4.2 Temporal flexibility and shared interpretation

Temporal flexibility changes when employees work and how much time they share with colleagues. Flexible schedules and asynchronous collaboration may support autonomy and focused work, but they may also weaken shared interpretation. Knowledge sharing is not only the movement of information between individuals. It also involves questioning, clarification, feedback, and collective sensemaking (Nonaka & Takeuchi, 1995; Weick, 1995). When employees work at different times, misunderstandings may remain unresolved for longer and shared meanings may develop more slowly.

This does not imply that asynchronous work is ineffective. Rather, temporal flexibility requires explicit routines for interpretation. Routine updates may be asynchronous, but ambiguous problems, innovation work, conflict resolution, and learning processes often require richer interaction. Without such routines, formal communication may still produce fragmented understanding.

Proposition 2: Temporal flexibility is more likely to produce knowledge fragmentation when asynchronous work is not supported by explicit feedback routines, decision documentation, and shared interpretation practices.

4.3 Organizational flexibility and digital knowledge visibility

Organizational flexibility concerns how work is coordinated through structures, routines, roles, and digital technologies. In flexible work settings, digital platforms often become the main infrastructure for communication, coordination, and knowledge access. They can strengthen knowledge visibility by making communication persistent, searchable, and accessible across locations, but may also hide context, interpretation, and tacit reasoning (Leonardi, 2014; Treem & Leonardi, 2013).

For this reason, digital visibility should not be equated with knowledge visibility. A repository may contain information without showing why it matters, how it should be interpreted, or who can explain its practical use. Flexible work may therefore strengthen knowledge processes only when digital tools are supported by shared norms for documentation, interpretation, access, and ownership (Majchrzak et al., 2013).

Proposition 3: Digital tools strengthen knowledge visibility in flexible work only when they are supported by shared norms for documentation, interpretation, access, and knowledge ownership.

4.4 Transactive memory systems and expertise coordination

An important element of organizational knowledge is knowing who knows what. Transactive memory systems allow teams to coordinate distributed expertise by developing shared awareness of members' knowledge domains (Lewis, 2003; Wegner, 1987). Flexible work may support this process by making expertise searchable across locations. At the same time, it may weaken expertise visibility because employees have fewer opportunities to observe problem-solving, informal referrals, and trust-based expertise recognition.

This creates expertise opacity: relevant knowledge may exist, but employees may not know where it resides or whom to consult. Because formal job titles and digital profiles rarely capture practical expertise, expertise visibility requires both technical infrastructure and social interaction.

Proposition 4: Flexible work is more likely to weaken transactive memory when expertise is not made visible through social interaction, transparent project histories, and accessible knowledge networks.

4.5 Knowledge governance and organizational outcomes

The framework proposes knowledge governance as the key condition linking flexible work to organizational knowledge outcomes. Knowledge governance refers to the intentional design of structures, technologies, routines, and managerial practices that shape knowledge creation, sharing, transfer, and use (Foss et al., 2010; Grandori, 2001). In flexible work contexts, governance becomes especially important because informal knowledge flows can no longer be assumed to arise from regular co-presence.

Knowledge governance should not be understood as excessive control. Its purpose is to make knowledge visible and usable by defining which interactions require synchronous discussion, how decisions are documented, where knowledge is stored, how newcomers access practice, and how expertise is made visible. When these mechanisms are absent, flexibility may increase individual autonomy while weakening collective knowledge visibility. When deliberately designed, flexible work may preserve autonomy while strengthening knowledge integration.

Proposition 5: Flexible work arrangements are more likely to support knowledge creation, sharing, retention, and renewal when organizations adopt deliberate knowledge governance mechanisms rather than relying on spontaneous knowledge flows.

5 Discussion and Implications

The framework developed in this paper contributes to the literature by reframing flexible work as a knowledge visibility problem. Existing research has mainly examined flexible work through individual outcomes such as autonomy, job satisfaction, wellbeing, productivity, and work-life balance (Allen et al., 2015; Gajendran & Harrison, 2007; Kossek & Lautsch, 2018). This paper extends that discussion by showing that flexible work also changes the organizational conditions under which knowledge is observed, interpreted, transferred, and renewed.

The first theoretical contribution is the separation of flexible work into spatial, temporal, and organizational flexibility. This distinction responds to the limitation of treating flexibility as a uniform arrangement. Spatial flexibility changes access to observation and informal learning; temporal flexibility changes feedback cycles and shared interpretation; organizational flexibility changes the technological and managerial infrastructure through which knowledge is coordinated. This provides a more precise basis for analysing flexible work outcomes and avoids deterministic claims that flexibility is either inherently beneficial or harmful.

The second contribution is the concept of knowledge visibility. Knowledge visibility refers not only to whether information is stored or accessible, but whether employees can understand its meaning, locate relevant expertise, and participate in the social processes through which knowledge becomes useful. This distinction matters because flexible work may increase documentation while still reducing shared understanding. Organizations may therefore appear knowledge-intensive at the level of digital records while becoming less knowledge-integrated at the level of practice. The broader theoretical contribution is the integration of knowledge management, digital work, and organizational learning perspectives into a single framework of flexible work as a socio-technical knowledge condition.

Managerially, organizations should not govern flexible work only through attendance rules or productivity indicators. They should also ask whether knowledge is visible, transferable, and interpretable. Knowledge governance should include critical knowledge mapping, decision logs, documentation standards, feedback routines, mentoring, guided newcomer participation, and communities of practice.

These implications are especially relevant for hybrid organizations. Flexible work may function well for experienced employees who already understand organizational routines, informal systems, and expertise

networks. The same arrangement may be more difficult for newcomers, cross-functional teams, or employees working on ambiguous and interdependent tasks. Flexible work policies should therefore be differentiated according to task interdependence, knowledge intensity, employee experience, and the degree of tacit knowledge involved.

The propositions provide a basis for future empirical research. Future studies could test whether spatial flexibility has stronger effects on tacit knowledge transfer in interdependent teams, whether decision logs and mentoring reduce knowledge fragmentation, and whether newcomer learning differs between formal onboarding and situated participation in hybrid work contexts.

6 Conclusions

This paper developed a conceptual framework explaining how flexible work arrangements reshape organizational knowledge visibility. The central argument is that flexible work should be understood not only as a matter of autonomy, productivity, or location choice, but also as a knowledge governance issue.

The paper contributes by distinguishing between spatial, temporal, and organizational flexibility and by explaining how each dimension affects different knowledge mechanisms. Spatial flexibility may reduce access to tacit learning and informal observation. Temporal flexibility may weaken shared interpretation when feedback and clarification are delayed. Organizational flexibility may strengthen digital traceability but may also produce tool-based fragmentation when information is dispersed without clear meaning or ownership. These effects depend on task interdependence, digital affordances, learning access, expertise visibility, and knowledge governance.

The proposed framework advances knowledge visibility as a bridge between flexible work and knowledge management research. Flexible work can support sustainable and knowledge-intensive organizations, but only when autonomy is matched with governance mechanisms that protect tacit learning, shared interpretation, expertise visibility, and newcomer participation.

7 Ethics declaration

This paper is conceptual and does not report empirical research involving human participants, personal data, or organizationally identifiable data. Therefore, ethical clearance was not required.

8 AI declaration

AI tools were used for language refinement, structural organization, and editorial support. The authors remained responsible for the conceptual development, theoretical framing, literature selection, interpretation, argumentation, and final manuscript.

References

Allen, T. D., Golden, T. D., & Shockley, K. M. (2015). How effective is telecommuting? Assessing the status of our scientific findings. *Psychological Science in the Public Interest*, *16*(2), 40-68. doi:10.1177/1529100615593273

Brown, J. S., & Duguid, P. (1991). Organizational learning and communities-of-practice: Toward a unified view of working, learning, and innovation. *Organization Science*, *2*(1), 40-57. doi:10.1287/orsc.2.1.40

Cheng, J., Sun, X., Zhong, Y., & Li, K. (2023). Flexible work arrangements and employees' knowledge sharing in post-pandemic era: The roles of workplace loneliness and task interdependence. *Behavioral Sciences*, *13*(2), 168. doi:10.3390/bs13020168

Eng, I., Aboagye, E., Bergsten, E., & Strömberg, A. (2025). Balancing employee flexibility and organizational performance: Implications for innovation, productivity, and company attractiveness in SMEs. *Frontiers in Psychology*, *16*, 1518284. doi:10.3389/fpsyg.2025.1518284

Foss, N. J., Husted, K., & Michailova, S. (2010). Governing knowledge sharing in organizations: Levels of analysis, governance mechanisms, and research directions. *Journal of Management Studies*, *47*(3), 455-482. doi:10.1111/j.1467-6486.2009.00870.x

Gajendran, R. S., & Harrison, D. A. (2007). The good, the bad, and the unknown about telecommuting: Meta-analysis of psychological mediators and individual consequences. *Journal of Applied Psychology*, *92*(6), 1524-1541. doi:10.1037/0021-9010.92.6.1524

Grandori, A. (2001). Neither hierarchy nor identity: Knowledge-governance mechanisms and the theory of the firm. *Journal of Management and Governance*, *5*(3-4), 381-399. doi:10.1023/A:1014045213456

Grant, R. M. (1996). Toward a knowledge-based theory of the firm. *Strategic Management Journal*, *17*(S2), 109-122. doi:10.1002/smj.4250171110

Heiman, B. A., & Nickerson, J. A. (2004). Empirical evidence regarding the tension between knowledge sharing and knowledge expropriation in collaborations. *Managerial and Decision Economics*, *25*(6-7), 401-420. doi:10.1002/mde.1198

Hesselbarth, Y., Alfes, K., & Festing, M. (2024). Understanding technology-driven work arrangements from a complexity perspective: A systematic literature review and an agenda for future research. *The International Journal of Human Resource Management*, *35*(5), 964-1006. doi:10.1080/09585192.2023.2298244

Kossek, E. E., & Lautsch, B. A. (2018). Work-life flexibility for whom? Occupational status and work-life inequality in upper, middle, and lower level jobs. *Academy of Management Annals*, *12*(1), 5-36. doi:10.5465/annals.2016.0059

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge, England: Cambridge University Press.

Leonardi, P. M. (2014). Social media, knowledge sharing, and innovation: Toward a theory of communication visibility. *Information Systems Research*, *25*(4), 796-816. doi:10.1287/isre.2014.0536

Lewis, K. (2003). Measuring transactive memory systems in the field: Scale development and validation. *Journal of Applied Psychology*, *88*(4), 587-604. doi:10.1037/0021-9010.88.4.587

Majchrzak, A., Faraj, S., Kane, G. C., & Azad, B. (2013). The contradictory influence of social media affordances on online communal knowledge sharing. *Journal of Computer-Mediated Communication*, *19*(1), 38-55. doi:10.1111/jcc4.12030

Nonaka, I. (1994). A dynamic theory of organizational knowledge creation. *Organization Science*, *5*(1), 14-37. doi:10.1287/orsc.5.1.14

Nonaka, I., & Takeuchi, H. (1995). *The knowledge-creating company: How Japanese companies create the dynamics of innovation*. New York, NY: Oxford University Press.

Polanyi, M. (1966). *The tacit dimension*. Garden City, NY: Doubleday.

Ren, Y., & Argote, L. (2011). Transactive memory systems 1985-2010: An integrative framework of key dimensions, antecedents, and consequences. *Academy of Management Annals*, *5*(1), 189-229. doi:10.5465/19416520.2011.590300

Spender, J. C. (1996). Making knowledge the basis of a dynamic theory of the firm. *Strategic Management Journal*, *17*(S2), 45-62. doi:10.1002/smj.4250171106

Torraco, R. J. (2005). Writing integrative literature reviews: Guidelines and examples. *Human Resource Development Review*, *4*(3), 356-367. doi:10.1177/1534484305278283

Treem, J. W., & Leonardi, P. M. (2013). Social media use in organizations: Exploring the affordances of visibility, editability, persistence, and association. *Annals of the International Communication Association*, *36*(1), 143-189. doi:10.1080/23808985.2013.11679130

Wegner, D. M. (1987). Transactive memory: A contemporary analysis of the group mind. In B. Mullen & G. R. Goethals (Eds.), *Theories of group behavior* (pp. 185-208). New York, NY: Springer.

Weick, K. E. (1995). *Sensemaking in organizations*. Thousand Oaks, CA: Sage.

Wenger, E. (1998). *Communities of practice: Learning, meaning, and identity*. Cambridge, England: Cambridge University Press.

Whetten, D. A. (1989). What constitutes a theoretical contribution? *Academy of Management Review*, *14*(4), 490-495. doi:10.5465/amr.1989.4308371

The Possible Effects of Digital Transformation on Social Enterprises and NGO

Maria Vittoria Prudentino, Paola Canestrini, Federica Bandini
University of Bologna. Bologna, Italy

Abstract

Existing research offers limited insight into how digital transformation unfolds in Social Enterprises (SEs). This study examines what drives SEs to engage in digital transformation, the facilitating and hindering factors, and work-related risks. Drawing on a theoretical framework integrating Dynamic Capabilities, Digital Social Innovation enablers and barriers, and Corporate Digital Responsibility, it adopts a qualitative multi-method case study based on a group interview and a focus group with managers from Italian SEs. The findings show that digital transformation is filtered through SEs' anthropocentric orientation. In the SEs studied, it is triggered by concrete organizational and social needs and leads to a reconfiguring process tailored to mission preservation. More specifically, digital transformation unfolds through processes of sensing, seizing, and reconfiguring, in which organizations identify unmet social and organizational needs, adopt technologies through gradual and context-sensitive approaches, and adapt internal structures, competences, and external partnerships over time. Digital transformation depends on a combination of human, organizational, and environmental factors, with hybrid competences, bottom-up experimentation, and ecosystem collaboration emerging as key enablers, while resource constraints, uneven digital literacy, and resistance to change pose obstacles. These barriers are not only structural but also reflect concerns related to service quality, inclusion, and mission coherence. The implications of digital transformation for work are actively filtered through mission-oriented values and responsibility-oriented practices, while digital technologies are predominantly framed as tools for mission augmentation rather than work substitution. In this respect, digital transformation in SEs appears to be governed through embedded responsibility-oriented logics, which shape how risks such as dehumanization and job displacement are anticipated and managed. Overall, the study contributes to research on digital transformation by showing how anthropocentric, mission-driven organizations interpret and enact digital change. It also suggests that other sectors may learn from SEs that maintaining a human-centered approach can help mitigate ethical and organizational risks associated with digital transformation.

Keywords: Social Economy, Digital Innovation, Dynamic Capabilities, Risks, Digital Humanism.

Paper type: Academic Research Paper

1 Introduction

Management studies identify Digital Transformation (DT) as one of the crucial and most pervasive challenges of recent decades (e.g. He et al., 2020; Brown et al., 2024). Moving beyond a technology-centric view, this literature emphasizes the organizational and human foundations of DT, including digital skills development, internal capability building (D'Angelo et al., 2024), and workforce arrangements (Dong et al., 2025). Social Enterprises (SEs) are as well increasingly engaging in DT (Yáñez-Valdés et al., 2023). Positioned at the intersection of market, public policies, and civil society (Nyssens, 2006), they respond to collective needs through socially oriented products, services, and models (Bandini, 2016; Defourny and Nyssen, 2017; Murray et al., 2010), thereby generating social innovation (Sottini et al., 2025). Despite this growing interest, studies on DT in SEs remain scarce. This is relevant because SEs are human-centered and mission-driven organizations (Vickers et al., 2025), where the implications of digital technologies for beneficiaries, workers, and communities are particularly salient.

While initial studies have begun to explore DT in SEs (Fan et al., 2025), knowledge of its organizational and ethical dimensions remains limited. Also, the ethical implications, particularly regarding labor, of DT within SEs remain underexplored compared to other sectors (Hanelt et al., 2021; Maiolini et al., 2025).

Hence, this study addresses the following research questions:

- *RQ1a: "What drives social enterprises to undertake digital transformation and how do they implement it at the organizational level?"*
- *RQ1b: "What are the enabling factors and obstacles SEs face in undertaking digital innovations?"*
- *RQ2: "How does digital transformation affect work in social enterprises, and how are potential risks of job loss and dehumanization addressed?"*

This research adopts three theoretical lenses. For RQ1a, the Dynamic Capabilities (DCs) (Teece, 1997; 2007). For RQ1b, frameworks on Digital Social Innovation (Buck et al., 2025; Buonocore et al., 2024). For RQ2, Corporate

Digital Responsibility (Lobschat et al., 2021). The study adopts a qualitative multi-method case study design (Yin, 2014), combining a group interview and a follow-up focus group with Italian SE managers.

The originality of this study lies in identifying how DCs unfold in SEs' DT, uncovering SE-specific enabling and hindering factors, and showing how their human-centered orientation limits risks of job loss and dehumanization. These findings suggest that the SE model may offer insights transferable beyond the sector.

2 Literature Review

Digital innovation refers to the use of digital technologies to develop new products, processes, and business models, reshaping organizational structures and capabilities and constituting DT (Nambisan et al., 2017; Hinings et al., 2018; Mergel et al., 2019; Hanelt et al., 2021; Nadkarni et al., 2021; Vial, 2019). In SEs, research has primarily examined digitalization through social innovation, highlighting how digital technologies support solutions to unmet social needs and enable collaboration across actors (Murray et al., 2010; Moulaert et al., 2013; Grassi and Toschi, 2021; Sottini et al., 2025; Qureshi et al., 2021; Yáñez-Valdés et al., 2023). SEs generate broader value-creation dynamics and rely on diverse technologies—from ICTs to AI—to support inclusion and service delivery (Gregori and Holzmann, 2020; Del Giudice et al., 2019; Bobsin et al., 2018; Sansone et al., 2023; Jong and Ganzaroli, 2023; Chen and Wang, 2024; Duong et al., 2025; Faludi, 2023; Turja et al., 2020). However, existing research focuses mainly on what technologies enable, offering limited insight into how transformation unfolds in anthropocentric organizations.

Prior studies identify enabling and constraining factors, including institutional conditions and digital infrastructure (Galindo-Martín, 2023; Li, 2023; Bhattacharyya, 2021; Vaño, 2025), as well as organizational challenges such as skill gaps and resource constraints (Sansone et al., 2023; Buonocore et al., 2024). Enabling factors span human, organizational, and environmental dimensions (Buck et al., 2025), while key challenges include technological change, stakeholder alignment, impact measurement, regulatory uncertainty, ethical dilemmas, and scalability (Buonocore et al., 2024).

DT also entails risks, including job displacement, dehumanization, and broader societal inequalities (Parasuraman and Riley, 1997; Khogali and Mekid, 2023; Cattaneo et al., 2025; Puntoni et al., 2020; Gursoy et al., 2025; Bircan and Özbilgin, 2025; Rothe et al., 2022). Digital technologies can simultaneously enable value creation and generate unintended consequences related to care services, data use, and AI (Čaić et al., 2018; Breidbach and Maglio, 2020; Zuboff, 2015; Madianou, 2021), raising concerns about exclusion, privacy, and misalignment with societal needs (Van Holstein et al., 2021; Zamani and Vannini, 2025; Shankar, 2024; Vaño, 2025). In this context, SEs' anthropocentric nature is crucial: they prioritize social value creation and relational well-being (Mair and Martí, 2006; Zahra et al., 2009; Smith et al., 2013; Santos, 2012; Dacin et al., 2010; Shepherd et al., 2023; Kummitha, 2025; Brieger et al., 2020), making responsible digital innovation a fundamental requirement.

3 Theoretical Framework

This study is structured around three research questions, each linked to a specific theoretical lens. For RQ1a we draw on Dynamic Capabilities (DCs) (Teece et al., 1997; Teece, 2007) conceptualize how organizations adapt through sensing, seizing, and reconfiguring, and have been applied to markets, public organizations, SEs, and DT (e.g. Piening, 2013; Vézina et al., 2018; Vallaster et al., 2019; Ince and Hahn, 2020; Bhardwaj et al., 2021; Pirotti et al., 2021) as well as to DT (Ellström et al., 2021; Chen et al., 2022; Kowalski et al., 2024; De La Torre and De La Vega, 2025) although their application to SE DT remains limited. For RQ1b, the study draws on Digital Social Innovation frameworks (Buck et al., 2025; Buonocore et al., 2024) to examine enabling and hindering factors of DT in SEs. Finally, for RQ2, Corporate Digital Responsibility (Lobschat et al., 2021) is used to analyze how organizations govern digital technologies through shared values, norms, and practices (Elliott and Copilah-Ali, 2024; Gursoy et al., 2025; Wirtz et al., 2022; Kummitha, 2025; Vaño, 2025; Gigauri et al., 2023; Zhou et al., 2025).

4 Methodology and Research Context

This study adopts a qualitative case study design (Yin, 2014) based on a group interview (Frey and Fontana, 1991) followed by a focus group (Morgan, 1996) with managers of Italian SEs. Participants were selected through purposive, criterion-based sampling to ensure information-rich cases (Saunders and Townsend, 2016; Plakoyiannaki and Budhwar, 2021). The sample included six participants from four SEs and, while not intended to be representative, provided sufficient depth for the exploratory aims of the study. The sequential use of the two methods enabled triangulation and cross-validation of findings (Jick, 1979; Eisenhardt, 1989; Christofi et al., 2024), as well as a more reflexive interpretation of participants' accounts (Robinson and Kerr, 2015). Data were

audio-recorded, transcribed verbatim, and analyzed using a hybrid inductive–deductive approach (Gioia et al., 2013; Strauss and Corbin, 2014; Saldaña, 2016). Initial open coding was followed by theory-driven analysis. Additional details on the research protocol, participant characteristics, interview guidelines, and coding procedures are available from the authors upon request.

Italy was selected as the empirical context due to its pioneering role in the development of SEs (Poledrini and Tortia, 2020; Defourny and Nyssens, 2010), particularly social cooperatives (Thomas, 2004). At the European level, SEs are increasingly recognized as key actors in the digital transition (European Commission, 2021; 2022; 2023), although empirical evidence confirms that Italian SEs still mainly rely on basic technologies, while advanced tools remain marginal (0.9%–9.8%) (ISTAT, 2021). Key barriers include limited financial resources, low digital culture, shortages of qualified personnel, and limited investment in innovation and training (ISTAT, 2021), in line with broader European trends (European Commission, 2023).

5 Findings and Discussion

DT in the SEs analyzed unfolds through sensing, seizing, and reconfiguring processes, consistent with prior research (Hinings et al., 2018; Hanelt et al., 2021; Nadkarni et al., 2021; Teece, 1997; 2007), but with a mission-oriented character. In the sensing phase, transformation is driven by concrete organizational and social needs rather than technological imperatives. Organizations assess internal inefficiencies, staff workload, and unmet social needs, extending beyond market-oriented opportunity recognition to include relational service quality and beneficiaries' well-being (Bhardwaj et al., 2021; Pirotti et al., 2021). In the seizing phase, technologies are introduced to improve efficiency and service quality by automating non-relational tasks and reallocating resources toward higher-value activities: "all administrative and bureaucratic aspects (…) can undergo processes of automation simplification, and digitalization(…) making it possible to redirect investments toward roles that generate higher added value". At the same time, they support users and staff, particularly in care contexts: "technology makes it possible to move beyond pharmacological approaches and improves the psychophysical well-being not only of service users but also of caregivers". Transformation is implemented gradually and through bottom-up processes rather than top-down imposition: "it is not technology that has to fit into the organization; rather, the organization must be able to open up and understand where technology can actually be adopted". Training and skill development emerge as central, alongside the attraction of specialized competencies (D'Angelo et al., 2024). The reconfiguring phase involves longer-term changes, including process revision, cultural adaptation, and competence development: "digital transformation is a long-term process (… that) affects organizational processes and people". It also includes strengthening external partnerships and, in some cases, governance transformation. Overall, these findings extend DCs applications to SEs (Piening, 2013; Vézina et al., 2018; Ince & Hahn, 2020; Fan et al., 2025), showing that capabilities are enacted through a dual orientation of adaptation and mission preservation.

DT is further shaped by human, organizational, and environmental conditions (Qureshi et al., 2021; Yáñez-Valdés et al., 2023; Buonocore et al., 2024; Buck et al., 2025). At the human level, hybrid competences combining technical, relational, and "translational" abilities enable alignment between technological and social needs (D'Angelo et al., 2024), while cultural resistance and limited digital literacy represent key barriers, as staff may "do not understand why a way of working that had worked until the day before needs to be changed". At the organizational level, mission alignment, identity, data use, and bottom-up experimentation act as enablers, whereas resource constraints significantly limit investments: "the first and main obstacle is the resource constraint, and some smaller organizations are not even able to overcome this initial barrier". Limited access to funding further reinforces these constraints. At the environmental level, collaboration with partners and beneficiary integration support innovation, while opportunity sensing and societal problem understanding enable meaningful applications. Data emerge as both enablers and challenges due to cybersecurity and management issues. These findings suggest that enabling and hindering factors are not merely structural but are organizationally interpreted, reflecting concerns related to service quality and mission coherence, while collaboration and experimentation extend prior work on SE value creation (Gregori & Holzmann, 2020; Del Giudice et al., 2019).

Finally, DT in SEs is grounded in shared values, norms, and practices oriented toward human agency. In high-relational contexts, technology is framed as support rather than substitution: "in the Social Economy, there is no risk of substitution because many activities have a strong relational component. Technology plays a role of augmentation rather than substitution". This aligns with the mission-driven nature of SEs (Ebrahim et al., 2014; Molecke & Pinkse, 2020; Fan et al., 2025; Vickers et al., 2025) and contrasts with broader concerns about job displacement and dehumanization (Parasuraman & Riley, 1997; Khogali & Mekid, 2023; Cattaneo et al., 2025; Zhou et al., 2025). At the level of norms, digital adoption is gradual and participatory, enabling workload

redistribution: “Digital technologies can be used to improve process efficiency by trying to free up resources from repetitive activities where people do not add value”. This is accompanied by training and attention to data protection. At the level of practices, these principles translate into concrete behaviors: “training is necessary to prevent a disconnect between staff and the organization” [A1], while “beyond technological facilitation, there is human facilitation”. Overall, these findings extend the CDR framework (Lobschat et al., 2021) by showing that responsibility is embedded ex ante in SEs, where DT is governed through participatory and mission-oriented practices.

6 Conclusion and Implication

This study examined how SEs undertake digital transformation, which factors enable or hinder this process, and how work-related risks are addressed in an anthropocentric organizational context. By combining a group interview and a focus group with Italian SE managers, it offers an exploratory but theoretically grounded account of digitalization as a process concerned not only with efficiency, but also with mission delivery and the protection of relational forms of work. The study highlights that digital competence in SEs should be understood holistically, as both technical proficiency and the organizational ability to mobilize skills, exercise judgment, and govern digital technologies consistently with social mission and stakeholder responsibilities (D’Angelo et al., 2024; Blanka et al., 2022; Fan et al., 2025; Lobschat et al., 2021).

The findings also suggest that higher education institutions can support SEs by fostering integrated digital skill development combining technical, ethical, and relational dimensions. From a policy perspective, they point to the need for partnerships among universities, public bodies, and SEs, dedicated financial support for digitalization, and monitoring of social impacts to ensure mission-consistent outcomes. More broadly, the SE context offers insights for organizations seeking digital transformation without compromising human agency, inclusion, and relational value.

The study has limitations: anthropocentric values may mitigate risks of dehumanization and job displacement, but such risks may not be absent in the long term; the small qualitative sample supports analytical insight rather than generalization; and managerial perspectives may underrepresent frontline workers, beneficiaries, and technology partners. Future research could adopt longitudinal, comparative, and multi-actor designs to examine how sensing, seizing, reconfiguring, and responsibility-oriented practices evolve across contexts, and how digital topics are integrated into SE training programs.

7 Acknowledgements

This research was supported by the European Union under the DigiSE5.0 project. The authors gratefully acknowledge this support.

Co-funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the Foundation for the Development of the Education System (FRSE). Neither the European Union nor FRSE can be held responsible for them.

8 Ethics Declaration

This study was conducted in accordance with the ethical and data management guidelines of the authors’ institution. All participants were informed about the purpose of the research, their voluntary participation, and the use of their contributions for research purposes. Written informed consent was obtained prior to data collection. All data were anonymized to ensure participants’ confidentiality.

9 AI Declaration

AI tools were selectively used in the writing and editing stages of this manuscript, primarily to support language refinement, text summarization, including the summarization of initial text drafts, and the organization of qualitative coding structures. The authors take full responsibility for the content of the manuscript and confirm that all AI-assisted outputs were carefully reviewed, verified, and revised to ensure their accuracy, validity, and consistency. All references and citations were independently checked to avoid inaccuracies. The authors acknowledge the limitations of AI tools, including potential bias and errors.

References

Bandini F. (2016). “Social enterprises” In Vecchi V., Balbo L., Brusoni M. and Caselli S. (Eds.). Principles and practice of impact investing, Greenleaf Pub

Bhardwaj, R., Srivastava, S., Taggar, R., & Bindra, S. (2021). Exploring micro-foundations of dynamic capabilities for social enterprises. Social Ent J, 18(3):451–469.

Bhattacharyya, S. S. (2021). Elucidating technology-based social entrepreneurs’ scale and scope perspectives. Int J of Org Analysis, 31(4):1001–1023.

Blanka, C., Krumay, B., & Rueckel, D. (2022). The interplay of digital transformation and employee competency. Tech Forecasting and Social Change, 178, 121575.

Bobsin, D., Petrini, M., & Pozzebon, M. (2018). The value of technology affordances to improve the management of nonprofit organizations. RAUSP Man J, 54(1):14–37.

Breidbach, C. F., & Maglio, P. (2020). Accountable algorithms? The ethical implications of data-driven business models. J of Service Man, 31(2):163–185.

Brieger, S. A., De Clercq, D., & Meynhardt, T. (2020). Doing good, feeling good? Entrepreneurs’ social value creation beliefs and Work-Related Well-Being. J of Business Ethics:172(4):707–725.

Brown, O., Davison, R. M., Decker, S., … (2024). Theory-driven perspectives on generative artificial intelligence in business and management. British J of Man, 35(1):3–23.

Buck, C., Heim, L., Körner-Wyrtki, K., Krombacher, A., & Röglinger, M. (2025). Making the most of digital social innovation. J of Business Research, 190:115215

Buonocore, F., Annosi, M. C., De Gennaro, D., & Riemma, F. (2024). Digital transformation and social change: Leadership strategies for responsible innovation. J of Eng and Tech Man, 74, 101843.

Čaić, M., Odekerken-Schröder, G., & Mahr, D. (2018). Service robots: value co-creation and co-destruction in elderly care networks. J of Service Man, 29(2):178–205.

Cattaneo, M. A., Gschwendt, C., & Wolter, S. C. (2025). How Scary is the Risk of Automation? Evidence from a Large Scale Survey Experiment. SSRN Electronic J.

Chen, J., & Wang, Z. (2024). How do artificial intelligence (AI) and big data (BD) technologies help social impact enterprises build legitimacy? J of Global Inf Man, 32(1):1–21.

Chen, Y., Li, J., & Zhang, J. (2022). Digitalisation, data-driven dynamic capabilities and responsible innovation: An empirical study of SMEs in China. Asia Pacific J of Man, 41(3):1211–1251.

Christofi, M., Hadjielias, E., Hughes, M., & Plakoyiannaki, E. (2024). Advancing research methodologies in management: Revisiting debates, setting new grounds for pluralism. British J of Man, 35(1):24–35.

D’angelo, S., Ghezzi, A., & Cavallo, A. (2024). Digital Skills Mobilization within Incumbent Organizations: The Agentic Role of Digital Champions. British J of Man, 35(2):594–612.

Dacin, P. A., Dacin, M. T., & Matear, M. (2010). Social entrepreneurship: why we don’t need a new theory and how we move forward from here. Academy of Man Perspectives, 24(3):37–57.

De La Torre, A., & De La Vega, I. (2025). Dynamic capabilities and digital innovation: pathways to competitive advantage through responsible innovation. J of Responsible Innovation, 12(1).

Defourny, J., & Nyssens, M. (2010). Conceptions of social enterprise and Social entrepreneurship in Europe and the United States: convergences and divergences. J of Social Entrepreneurship, 1(1):32–53.

Defourny, J., & Nyssens, M. (2017). Fundamentals for an international typology of Social Enterprise Models. Voluntas, 28(6):2469–2497.

Del Giudice, M., Garcia-Perez, A., Scuotto, V., & Orlando, B. (2019). Are social enterprises Tech innovative? A quantitative analysis on social entrepreneurs in emerging countries. Tech Forecasting and Social Change, 148, 119704.

Dong, M., Choi, Y., Ma, P., Cai, H., & Cui, L. (2024). Employee ownership and firm digital transformation. British J of Man, 36(2):511–528.

Duong, C. D., Nguyen, T. H., … (2025). Exploring the role of generative artificial intelligence (ChatGPT) adoption in digital social entrepreneurship. Social Enterprise J.

Ebrahim, A., Battilana, J., & Mair, J. (2014). The governance of social enterprises: Mission drift and accountability challenges in hybrid organizations. Research in Org Behavior, 34, 81–100.

Eisenhardt, K. M. (1989). Building theories from case study research. Academy of Man Review, 14(4):532–550.

Elliott, K., & Copilah-Ali, J. (2024). Implementing corporate digital responsibility (CDR). Org Dynamics, 53(2):101040.

Ellström, D., Holtström, J., Berg, E., & Josefsson, C. (2021). Dynamic capabilities for digital transformation. J of Strategy and Man, 15(2):272–286.

EU Regulation No. 1689 (2024) of the European Parliament and of the Council of 13 June 2024 laying down harmonized rules on artificial intelligence (Artificial Intelligence Act).

European Commission. (2021). Building an economy that works for people: An action plan for the social economy. Publications Office of the European Union. European Commission.

European Commission. (2022). Transition Pathway for Proximity and Social SOI: 10.2873/44306 Economy.

European Commission. (2023). Annex: Italy. In 2030 Digital Decade – Report on the state of the Digital Decade 2023.

Faludi, J. (2023). How to create social value through digital Social Innovation?. J of Social Entrepreneurship, 14(1):73–90

Fan, X., Liu, Y., Chen, A., & Shou, Y. (2025). How to digitalize social enterprises. British J of Man, 36(4):1635–1652.

Frey, J. H., & Fontana, A. (1991). The group interview in social research. The Social Science J, 28(2):175–187.

Galindo-Martín, M., Castaño-Martínez, M., & Méndez-Picazo, M. (2023). Digitalization, entrepreneurship and competitiveness. Rev of Man Science, 17(5):1809–1826.
Gigauri, I., Apostu, S., Popescu, C. (2023) Digital Transformation: Threats and Opportunities for Social Entrepreneurship, In Akkaya, B. & Tabak, A. (Eds.), Two Faces of Digital Transformation: Tech Opportunities versus Social Threats
Gioia, D. A., Corley, K. G., & Hamilton, A. L. (2013). Seeking qualitative rigor in inductive research, 16(1):15–31.
Glasbeek, L. (2024). Bricolage and its Strategic Connotations. British J of Man, 36(3):1003–1022.
Grassi, E., Toschi, L. (2021). A systematic literature review of technology social ventures: state of the art and directions for future research at the Micro-, Meso-and Macro- Level. J of Social Entrepreneurship, 15(3):1–33.
Gregori, P., & Holzmann, P. (2020). Digital sustainable entrepreneurship: A business model perspective on embedding digital technologies for social and environmental value creation. J of Cleaner Production, 272, 122817.
Gursoy, D., Başer, G., & Chi, C. G. (2025). Corporate digital responsibility: navigating ethical, societal, and environmental challenges in the digital age and exploring future research directions. J of Hospitality Marketing & Man, 34(3):305–324.
Hanelt, A., Bohnsack, R., Marz, D. and Antunes Marante, C. (2021). A Systematic Review of the Literature on Digital Transformation. J of Man Studies, 58(5):1159-1197.
He, Q., Meadows, M., Angwin, D., Gomes, E., & Child, J. (2020). Strategic Alliance Research in the Era of Digital Transformation. British J of Man, 31(3):589–617.
Hinings, B., Gegenhuber, T., & Greenwood, R. (2018). Digital innovation and transformation. Inf and Organization, 28(1):52–61.
ISTAT. (2021). Permanent census of non-profit institutions.
Jick, T. D. (1979). Mixing qualitative and quantitative methods. Adm Science Q, 24(4):602–611.
Jong, C. L. K., & Ganzaroli, A. (2023). Managing Digital Transformation for Social Good in Non-Profit Organizations. Voluntas, 35(3):417–427.
Khogali, H. O., & Mekid, S. (2023). The blended future of automation and AI. Tech in Society, 73, 102232.
Kowalski, M., Bernardes, R. C., Gomes, L., & Borini, F. M. (2024). Microfoundations of dynamic capabilities for digital transformation. European J of Innov Man, 28(8):3717–3746.
Kummitha, R. K. R., Honig, B., & Urbano, D. (2025). Social Entrepreneurship: a Well-Being based approach. J of Business Ethics, 200(3):557–597.
Li, B. (2023). Effects of digital economy on social entrepreneurship. Managerial and Decision Economics, 44(8):4248–4261.
Lobschat, L., Mueller, B., Eggers, F., Bradimarte, L., Diefenbach, S., Kroschke, M., Wirtz, J. (2021), Corporate Digital Responsibility, J of Business Research, 122, 878-888.
Madianou, M. (2021). Nonhuman humanitarianism: when "AI for good" can be harmful. Inf Communication & Society, 24(6):850–868.
Maiolini R., Capo F. and Venturi P. (2025). Innovating for social good: digital transformation and crisis-based entrepreneurial opportunities. J of Small Business and Enterprise Dev, 32(4):981–1000
Mair, J., & Martí, I. (2006). Social entrepreneurship research. J of World Business, 41(1):36–44.
Mergel, I., Edelmann, N., & Haug, N. (2019). Defining digital transformation. Government Inf Q, 36(4):101385
Molecke, G., & Pinkse, J. (2020). Justifying social impact as a form of impression management. British J of Man, 31(2):387–402
Morgan, D. L. (1996). Focus groups. Annual Review of Sociology, 22(1):129–152.
Moulaert, F., MacCallum, D., & Hillier, J. (2013). Social innovation: intuition, precept, concept, theory and practice. In Edward Elgar Publishing eBooks.
Murray, R., Caulier-Grice, J., & Mulgan, G. (2010). The open book of social innovation. The Young Foundation/NESTA.
Nadkarni, S., & Prügl, R. (2021). Digital transformation: a review, synthesis and opportunities for future research. Man Review Q, 71(2):233–341.
Nambisan, S., Lyytinen, K., Majchrzak, A., & Song, M. (2017). Digital Innovation Management: Reinventing innovation management research in a digital world. MIS Q, 41(1):223–238
Nyssens, M. (Ed.). (2006). Social enterprise: At the crossroads of market, public policies and civil society. Routledge.
Parasuraman, R., & Riley, V. (1997). Humans and automation. J of the Human Factors and Ergonomics Society, 39(2):230–253.
Piening, E. P. (2013). Dynamic capabilities in public organizations. PMR, 15(2):209–245.
Pirotti, T. M. C., Bitencourt, C. C., Faccin, K., & Kretschmer, C. (2021). The Process of Social Innovation Scalability. J of Innovation Man, 9(2):21–38.
Plakoyiannaki, E., & Budhwar, P. (2021). From convention to alternatives: Rethinking qualitative research in management scholarship. British J of Man, 32(1):3–6.
Poledrini, S., & Tortia, E. C. (2020). Social Enterprises: evolution of the organizational model and application to the Italian case. Entrepreneurship Research J, 10(4).
Puntoni, S., Reczek, R. W., Giesler, M., & Botti, S. (2020). Consumers and Artificial intelligence. J of Marketing, 85(1):131–151
Qureshi, I., Pan, S. L., & Zheng, Y. (2021). Digital social innovation: An overview and research framework. Inf Systems J, 31(5):647–671
Robinson, S., & Kerr, R. (2015). Reflexive conversations: Constructing hermeneutic designs for qualitative management research. British J of Man, 26(4):777–790.
Rothe, F., Van Audenhove, L., & Loisen, J. (2022). Digital development, inequalities & the Sustainable Development Goals. Inf Technology for Development, 29(1):9–26.

Sahasranamam, S., Lall, S., Nicolopoulou, K., & Shaw, E. (2024). Founding Team Entrepreneurial experience, external financing and social enterprise performance. British J of Man, 35(1):519–536.
Saldaña, J. (2016). The Coding Manual for Qualitative Researchers (3rd ed.). London: SAGE Publications.
Sansone, G., Santalucia, F., Viglialoro, D., & Landoni, P. (2023). Blockchain for social good and stakeholder engagement: Evidence from a case study. Corporate Social Resp and Env Man, 30(5):2182–2193.
Santos, F. M. (2012). A positive theory of social entrepreneurship. J of Business Ethics, 111(3):335–351.
Saunders, M. N. K., & Townsend, K. (2016). Reporting and justifying the number of interview participants in organisation and workplace research. British J of Man, 27(4):836–852.
Shankar, V. (2024). Managing the Twin Faces of AI: A commentary on "Is AI changing the world for better or worse?" J of Macromarketing.
Shepherd, D. A., Seyb, S., & Williams, T. A. (2023). Empathy-driven entrepreneurial action: Well-being outcomes for entrepreneurs and target beneficiaries. J of Business Venturing, 38(2):106290.
Smith, W. K., Gonin, M., & Besharov, M. L. (2013). Managing Social-Business Tensions: A Review and Research Agenda for Social Enterprise. Business Ethics Q, 23(3):407–442.
Sottini A.C.M., Zupic I. and Giudici A. (2025). Social entrepreneurship and social innovation. European Man J, 43(2):182-193;
Strauss, A., & Corbin, J. (2014). Basics of qualitative research, 4th ed., Sage.
Teece, D. J. (2007). Explicating dynamic capabilities. Strategic Man J, 28(13):1319–1350.
Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic capabilities and strategic management. Strategic Man J, 18(7):509–533. https://doi.org/10.1002/(sici)1097-0266(199708)18:7
Thomas, A. (2004). The rise of social cooperatives in Italy. Voluntas, 15(3):243–263
Turja, T., Aaltonen, I., Taipale, S., & Oksanen, A. (2020). Robot acceptance model for care. Inf & Man, 57(5):103220
Vallaster, C., Maon, F., Lindgreen, A., & Vanhamme, J. (2019). Serving Multiple Masters: The role of micro-foundations of dynamic capabilities in addressing tensions in for-profit hybrid organizations. Org Studies, 42(6):911–947.
Van Holstein, E., Wiesel, I., Bigby, C., & Gleeson, B. (2021). People with intellectual disability and the digitization of services. Geoforum, 119:133–142.
Vañó, M. J. V. (2025). Social entrepreneurship: challenges, risks, and opportunities. Int Entrepreneurship and Man J, 21(1).
Vézina, M., Selma, M. B., & Malo, M. C. (2018). Exploring the social innovation process in a large market based social enterprise. Man Decision, 57(6):1399–1414.
Vial, G. (2019). Understanding digital transformation: A review and a research agenda. The J of Strategic Inf Systems, 28(2):118–144.
Vickers, I., Lyon, F., & Sepulveda, L. (2025). Collective capabilities for organizational democracy. British J of Man, 36(1):240–254.
Wirtz, J., Kunz, W. H., Hartley, N., & Tarbit, J. (2022). Corporate digital responsibility in service firms and their ecosystems. J of Service Research, 26(2):173–190
Yáñez-Valdés, C., Guerrero, M., Barros-Celume, S., & Ibáñez, M. J. (2023). Winds of change due to global lockdowns: Refreshing digital social entrepreneurship research paradigm. Tech Forecasting and Social Change, 190, 122454.
Yin, R. K. (2014). Case study research: Design and methods (5th ed.). SAGE.
Zahra, S. A., Gedajlovic, E., Neubaum, D. O., & Shulman, J. M. (2009). A typology of social entrepreneurs. J of Business Venturing, 24(5):519–532.
Zamani, E. D., & Vannini, S. (2025). Digital policy narratives: addressing grand challenges or exacerbating digital inequalities? Inf Technology for Development, 1–20.
Zhou, Q., Yu, H., Chen, M., Snell, M., & Adams, K. (2025). Managing Tech Innovation for Social Good. Strategic Change.
Zuboff, S. (2015). Big other: Surveillance Capitalism and the Prospects of an Information Civilization. J of Inf Technology, 30(1):75–89

Integrating Knowledge Management and Relational Capital to Improve Elderly Care: An Interventionist Approach

Francesca Dal Mas[1,2], Ludovica Schiavone[3], Giuseppe Pezzotti[4], Alberto Cavazza[2,3]

[1]Cà Foscari University of Venice, Venice, Italy
[2]SAN University, Lodz, Poland
[3]University of Milan-Bicocca, Milan, Italy
[4]Kansai Medical University, Hirakata, Japan

Abstract

The aging population is a major systemic challenge of the 21st century, reshaping healthcare demand, organizational structures, and knowledge dynamics. In highly aged countries like Japan, demographic pressures expose fragmentation in care pathways and the limits of purely structural or technological solutions. This paper sees Knowledge Management (KM) and Relational Capital (RC) are strategic assets that enable healthcare organizations to turn aging from a liability into a catalyst for learning, coordination, and innovation.

The study examines the interplay between KM and RC within a university-hospital ecosystem, specifically analyzing Kansai Medical University (KMU) and its affiliated hospital in Japan. It employs an Interventionist Research approach, using six months of participant observation, institutional documentation, and interviews to deepen the integration of knowledge translation into routine organizational processes.

The results indicate that KMU functions as a relational nexus where clinical, organizational, and community knowledge converge. Formal procedures, such as interprofessional rounds, digital repositories, and structured training, enhance knowledge generation and dissemination, whereas informal, trust-based networks promote the exchange of tacit knowledge. Collectively, these processes strengthen coordination, resilience, and continuity of care in response to population shifts.

The study emphasizes that traditional Japanese caregiving concepts are culturally ingrained in senior care, yet are increasingly implemented through modern knowledge and relational frameworks. In this setting, RC links hospital services with community-based care, facilitating integrated models in which clinical transitions are processes of knowledge translation supported by relational continuity.

Digital technologies, such as telemedicine, digital platforms, and AI-driven tools, augment these dynamics by improving service delivery and enabling trust-based interactions in virtual settings. Nonetheless, the research shows that technology is effective only when integrated within robust relational and knowledge frameworks, functioning as an enabler rather than a replacement.

The university serves as a pivotal coordinator of the healthcare knowledge ecosystem, integrating education, research, and care delivery while ensuring the stability of knowledge flows across many stakeholders. The study proposes the relational hub as an empirically grounded configuration that renders KM and RC observable and actionable within healthcare ecosystems. It implies that effective responses to aging require organized knowledge-sharing systems, intentional development of research collaborations beyond hospital confines, and a university-led coordination role to promote innovation.

Keywords: Knowledge Management; Relational Capital; Aging Population; Healthcare Ecosystems; Interventionist Research

Paper type: Academic Research Paper

1 Introduction

The aging population poses a significant systemic challenge for healthcare systems in the twenty-first century. Rising life expectancy, chronic illnesses, and multimorbidity heighten the need for long-term, integrated care while revealing organizational fragmentation, workforce shortages, and coordination challenges. Japan is a particularly salient context because its demographic profile exposes the tensions between rising care demands and the sustainability of healthcare provision earlier than in many other nations (Bratianu et al., 2024; Sudo et al., 2018).

The primary challenge lies in both clinical capacity and the ability to generate, disseminate, and use knowledge across professional and organizational boundaries. KM is crucial for converting fragmented information into actionable insights, enhancing decision-making, and fostering innovation in complex care environments (Abubakar et al., 2019; El Morr and Subercaze, 2010; Nonaka et al., 1996; Tsoukas and Vladimirou,

2001). Healthcare organizations often struggle to integrate specialized expertise, standardized processes, and patient data in a cohesive way. This is especially vital in geriatric care, where continuity depends on collaboration among hospitals, primary care providers, rehabilitation agencies, families, and communities.

RM enhances KM by fostering trust-based relationships, institutional collaborations, and cooperative practices that facilitate knowledge dissemination and coordinated efforts. The dynamics of senior care in Japan are shaped by cultural values of self-help (Ji-jo), mutual aid (Go-jo), social solidarity (Kyo-jo), and government care (Ko-jo), which continue to inform community-based care arrangements (Hatano et al., 2017; Sudo et al., 2018). At the same time, digital technologies are expanding opportunities for knowledge dissemination, telemedicine, and remote coordination; yet their effectiveness depends on integration within comprehensive knowledge ecosystems (Bagnoli et al., 2019; Biancone et al., 2021; Ito et al., 2017; Song, 2024).

This research examines how KM and resource coordination jointly enhance healthcare responses to aging, using KMU and its affiliated hospital as a case study. The university is viewed not merely as an educational entity but as a facilitator of a multifaceted knowledge ecosystem that integrates care, research, training, and community participation (Granstrand and Holgersson, 2020; Massaro et al., 2022; Secundo et al., 2018; Thomas et al., 2021).

This study examines the following research question:

RQ: Which knowledge management and resource coordination solutions are most effective in addressing the issues posed by an aging population within a healthcare ecosystem?

2 Literature Review

Healthcare is inherently knowledge-intensive. Clinical practice relies on established evidence, professional knowledge, experiential learning, and ongoing assessment of patient situations. Consequently, KM in healthcare transcends mere information storage; it encompasses the processes of knowledge creation, sharing, translation, and application to enhance care and organizational performance (Abubakar et al., 2019; El Morr and Subercaze, 2010; Nonaka et al., 1996; Tsoukas and Vladimirou, 2001). The significance of KM in aged care increases because patients frequently have multiple comorbidities, longer care trajectories, and a need for coordinated interventions over time. Knowledge translation is essential to connect evidence, professional competencies, and practical care practices (Bratianu et al., 2024; Dal Mas et al., 2020).

The literature indicates that knowledge processes in healthcare are seldom efficient when conducted in isolation. They rely on relationships that facilitate collaboration among various stakeholders, especially physicians, nurses, therapists, administrators, patients, and caregivers. RC denotes the value produced by trust-based relationships and the networks that facilitate the transfer of information across organizational borders. In healthcare ecosystems, RC enhances co-production, fosters coordination, and advances patient-centered integration (Cobianchi et al., 2022; Hatano et al., 2017; Zamboni et al., 2024). Relational structures are integral to KM; they are essential for the circulation of tacit information and the appropriate interpretation of codified knowledge.

Digital transformation adds a new dimension to these processes. Telemedicine, e-health platforms, remote monitoring, and AI-assisted technologies can expedite diagnosis, enhance accessibility, and expand the data available for clinical decision-making (Bagnoli et al., 2019; Biancone et al., 2021; Ito et al., 2017). Nonetheless, large volumes of data do not inherently become valuable knowledge. Their value depends on organizational practices that enable interpretation, dissemination, and appropriate use (Obenshain, 2004; Song, 2024; Wickramasinghe et al., 2009). From a KM perspective, digitization facilitates knowledge transfer rather than replacing human discernment and relationship continuity.

An expanding body of research highlights the significance of universities within knowledge ecosystems. Universities produce scientific knowledge, educate professionals, and increasingly serve as intermediaries linking many stakeholders in innovation processes (Granstrand and Holgersson, 2020; Massaro et al., 2022; Secundo et al., 2018; Thomas et al., 2021). In healthcare, the orchestration function is particularly pertinent given demographic aging, which necessitates ongoing education, interdisciplinary training, and aligning hospital procedures with community needs. Nonetheless, empirical research demonstrating the interaction between KM and resource capabilities in daily organizational routines to tackle aging-related difficulties remains scarce. This paper addresses that gap through an interventional case study of KMU.

3 Methodology

The study adopts interventionist research (IVR), a qualitative approach that combines theoretical ambition with close engagement in organizational settings. IVR is particularly appropriate when the objective is to understand how knowledge processes unfold in practice and how actors make sense of organizational issues

while responding to them in real time (Lukka and Wouters, 2022). The approach also allows the researcher to alternate between an insider position, needed to observe routines and interactions, and an outsider perspective, useful for linking empirical observations to broader conceptual debates.

The empirical setting is the KMU complex in Hirakata, Osaka Prefecture, comprising the university campus and the affiliated university hospital. KMU was selected because it combines advanced medical specialization, strong links with the local healthcare network, and a clear educational mission. Its institutional identity, grounded in the principle of Jijinshinkyo (benevolence, compassion, and empathy) is reflected in a broader medium-term vision that connects education, research, healthcare delivery, rehabilitation, business management, and social contribution. This makes KMU a particularly suitable setting for examining how knowledge and relationships are orchestrated in response to demographic aging.

Data collection lasted six months and involved three complementary sources. First, participant observation enabled the researcher to follow meetings, lectures, inter-professional interactions, and everyday practices across the KMU ecosystem. Second, institutional documentation and public materials were analyzed to reconstruct organizational structures, strategic priorities, and partnership development. Third, semi-structured interviews with key stakeholders were conducted to validate emerging interpretations and to explore how actors understood the roles of KM, RC, and digital transformation in elderly care. Data analysis followed an iterative logic, moving back and forth between field evidence and the KM literature to identify recurring patterns of knowledge creation, knowledge translation, coordination, and ecosystem orchestration.

4 Findings

The analysis shows that KMU functions as a relational hub where clinical, organizational, and community knowledge converge. The rising incidence of chronic and age-related conditions, including stroke, cardiovascular diseases, and neurodegenerative disorders, pressures not only clinical services but also the ability to circulate knowledge across departments. In this environment, elderly care depends on how quickly lessons learned in one unit can inform decisions in another. Knowledge sharing is therefore embedded in routines such as interdepartmental meetings, case discussions, and coordinated transitions between acute care and follow-up services.

Formal KM mechanisms play a central role in this process. Interprofessional rounds, shared repositories, and regular training activities create spaces where codified knowledge can be updated and disseminated, while also allowing professionals to collectively interpret the implications of new cases. These routines strengthen the hospital's absorptive capacity by making it possible to recognize relevant knowledge, integrate it into practice, and adapt it to local needs. At the same time, tacit knowledge remains essential. Informal exchanges among doctors, nurses, pharmacists, therapists, and support staff transmit experiential insights that cannot be fully captured in protocols but are fundamental for elderly patients whose conditions evolve over time and often require nuanced judgment.

RC makes these knowledge processes durable across institutional boundaries. Field evidence indicates that the traditional Japanese principles of Ji-jo, Go-jo, Kyo-jo and Ko-jo are not only cultural values but also organizing principles that sustain continuity of care. At KMU, the transfer of elderly patients from the acute hospital to local facilities, rehabilitation services and home-based care is treated as a knowledge translation process: information, clinical reasoning and responsibility are transferred together. Trust among KMU, smaller hospitals, clinics and community actors enable this circulation of knowledge and reduces the risk of fragmentation. Home care, visiting nursing and rehabilitation services therefore appear as extensions of the hospital's knowledge network rather than as separate domains.

Digital technologies expand, but do not replace, these relational and knowledge infrastructures. KMU is investing in telemedicine, AI-assisted diagnostics, remote monitoring and mobile hospital services able to reach peripheral areas. These tools extend the visibility of patient conditions and create new opportunities for coordination, especially for older adults with mobility constraints or limited family support. Yet their contribution is meaningful only when data are interpreted collectively, embedded in clinical routines and linked to trusted channels of communication. In this sense, digital transformation enhances virtual relational capital by connecting patients and professionals at a distance, but it remains dependent on the human and organizational conditions of KM.

The university dimension is decisive in stabilizing these mechanisms over time. KMU does not simply provide clinical services; it educates students, residents, and young doctors, recruits many of its own graduates, and organizes continuous learning through research centers, simulation activities, and interdisciplinary training. This educational role helps preserve knowledge continuity despite workforce turnover and regulatory constraints. It also supports the development of shared geriatric competencies across professions, which is essential in a

context where elderly care cannot be managed through narrow specialization alone. Community engagement initiatives and international collaborations further expand the ecosystem, positioning KMU as a platform through which local and global knowledge can be connected.

5 Discussion

This paper analyzed the role of KM and RC in addressing demographic aging in the healthcare sector, using the case study of KMU in Japan. The findings indicate that successful aged care depends on the ability to synchronize the generation, sharing, and translation of knowledge among hospital departments, community partners, and digital platforms. RC facilitates these processes by maintaining trust-based teamwork and continuity of care, while the university functions as an orchestrator that aligns education, research, and service delivery.

This research enhances KM literature by emphasizing the strategic importance of knowledge infrastructures in senior care and by defining the relational hub as the organizational framework through which KM and RC are operationalized. It also offers practical insights for healthcare managers and policymakers: demographic aging should be addressed through system-wide knowledge initiatives rather than discrete technology or structural interventions.

The study has limitations. It is based on a single case and employs a qualitative, interpretive framework. Future research may analyze various healthcare ecosystems, assess the long-term effects of specific KM strategies, and investigate how digital tools transform knowledge governance in aged care across diverse institutional contexts.

6 Ethics Declaration

Under the applicable institutional requirements for this study, formal ethical clearance was not required. The research was based on participant observation in organizational settings, analysis of institutional documentation and semi-structured interviews focused on professional roles and organizational practices. Participation was voluntary, participants were informed about the purpose of the study, and informed consent was obtained before data collection.

7 AI Declaration

AI tools were solely used to check language consistency, and not to generate new content.

References

Abubakar, A.M., Elrehail, H., Alatailat, M.A. and Elçi, A. (2019), “Knowledge management, decision-making style and organizational performance”, Journal of Innovation and Knowledge, Vol. 4 No. 2, pp. 104-114.

Bagnoli, C., Dal Mas, F. and Massaro, M. (2019), “The 4th Industrial Revolution: business models and evidence from the field”, International Journal of E-Services and Mobile Applications, Vol. 11 No. 3, pp. 34-47.

Biancone, P., Secinaro, S., Marseglia, R. and Calandra, D. (2021), “E-health for the future. Managerial perspectives using a multiple case study approach”, Technovation, Vol. 120, p. 102406, doi: 10.1016/j.technovation.2021.102406.

Bratianu, C., Garcia-Perez, A., Dal Mas, F. and Bedford, D. (2024), “Knowledge translation in healthcare”, Knowledge Translation, Emerald Publishing Limited, pp. 155-167, doi: 10.1108/978-1-80382-889-320241010.

Cobianchi, L., Dal Mas, F., Massaro, M., Bednarova, R., Biancuzzi, H., Filisetti, C., Barcellini, A., Orlandi, E., Miceli, L. and Angelos, P. (2022), “Hand in hand: a multistakeholder approach for co-production of surgical care”, The American Journal of Surgery, Vol. 223 No. 1, pp. 214-215, doi: 10.1016/j.amjsurg.2021.07.053.

Dal Mas, F., Biancuzzi, H., Massaro, M. and Miceli, L. (2020a), “Adopting a knowledge translation approach in healthcare co-production. A case study”, Management Decision, Vol. 58 No. 9, pp. 1841-1862, doi: 10.1108/MD-10-2019-1444.

Dal Mas, F., Garcia-Perez, A., Sousa, M.J., da Costa, R.L. and Cobianchi, L. (2020b), “Knowledge translation in the healthcare sector. A structured literature review”, Electronic Journal of Knowledge Management, Vol. 18 No. 3, pp. 198-211.

El Morr, C. and Subercaze, J. (2010), “Knowledge management in healthcare”, Handbook of Research on Developments in E-Health and Telemedicine: Technological and Social Perspectives, IGI Global Scientific Publishing, pp. 490-510.

Granstrand, O. and Holgersson, M. (2020), “Innovation ecosystems: a conceptual review and a new definition”, Technovation, Vols 90-91, p. 102098, doi: 10.1016/j.technovation.2019.102098.

Hatano, Y., Matsumoto, M., Okita, M., Inoue, K., Takeuchi, K., Tsutsui, T., Nishimura, S., et al. (2017), “The vanguard of community based integrated care in Japan: the effect of a rural town on national policy”, International Journal of Integrated Care, Vol. 17 No. 2, p. 2.

Ito, J., Edirippulige, S., Aono, T. and Armfield, N.R. (2017), “The use of telemedicine for delivering healthcare in Japan: systematic review of literature published in Japanese and English languages”, Journal of Telemedicine and Telecare, Vol. 23 No. 10, pp. 828-834, doi: 10.1177/1357633X17732801.

Lukka, K. and Wouters, M. (2022), “Towards interventionist research with theoretical ambition”, Management Accounting Research, Vol. 55, p. 100783, doi: 10.1016/j.mar.2022.100783.

Massaro, M., Dal Mas, F. and Bagnoli, C. (2022), “Academic moral entrepreneurship and knowledge translation to turn crises into opportunities: the case of VeniSIA”, IEEE Transactions on Engineering Management, Vol. 71, pp. 13446-13456, doi: 10.1109/TEM.2022.3188035.

Nonaka, I., Takeuchi, H. and Umemoto, K. (1996), “A theory of organizational knowledge creation”, International Journal of Technology Management, Vol. 11 Nos 7-8, pp. 833-845.

Obenshain, M.K. (2004), “Application of data mining techniques to healthcare data”, Infection Control and Hospital Epidemiology, Vol. 25 No. 8, pp. 690-695.

Secundo, G., Massaro, M., Dumay, J.C. and Bagnoli, C. (2018), “Intellectual capital management in the fourth stage of IC research: a critical case study in university settings”, Journal of Intellectual Capital, Vol. 19 No. 1, pp. 157-177, doi: 10.1108/JIC-11-2016-0113.

Song, H. (2024), The Role of Information and Communication Technologies in Knowledge Management: From Enabler to Facilitator, RMIT University.

Sudo, K., Kobayashi, J., Noda, S., Fukuda, Y. and Takahashi, K. (2018), “Japan’s healthcare policy for the elderly through the concepts of self-help (Ji-jo), mutual aid (Go-jo), social solidarity care (Kyo-jo), and governmental care (Ko-jo)”, BioScience Trends, Vol. 12 No. 1, pp. 7-11.

Thomas, E., Faccin, K. and Asheim, B.T. (2021), “Universities as orchestrators of the development of regional innovation ecosystems in emerging economies”, Growth and Change, Vol. 52 No. 2, pp. 770-789.

Tsoukas, H. and Vladimirou, E. (2001), “What is organizational knowledge?”, Journal of Management Studies, Vol. 38 No. 7, pp. 973-993, doi: 10.1111/1467-6486.00268.

Wickramasinghe, N., Bali, R.K., Lehaney, B., Schaffer, J. and Gibbons, M.C. (2009), Healthcare Knowledge Management Primer, Routledge.

Zamboni, F., Paoloni, P., Cavazza, A. and Dal Mas, F. (2024), “Virtual relational capital for business development: a case study”, VINE Journal of Information and Knowledge Management Systems, Vol. 56 No. 1, pp. 1-18, doi: 10.1108/VJIKMS-06-2024-0226.

Beyond Facilitation: Ethical and Behavioral Tensions in Technology-Enabled Knowledge Management

Emilia Romeo[1], Valerio Giampaola[2], Giuliano Marolla[3], Angelo Rosa[3], Alberto Ferraris[4]
[1]San Raffaele Roma University, Rome, Italy
[2]eCampus University, Novedrate (CO), Italy
[3]LUM University Giuseppe Degennaro, Casamassima (BA), Italy
[4]University of Turin, Turin (TO), Italy

Abstract

This study examines the ethical and behavioural implications of technology-enabled knowledge management (KM) systems, challenging the dominant view of digital technologies as neutral tools that merely facilitate knowledge processing. Instead, it conceptualises KM systems as sociotechnical infrastructures that actively shape how knowledge is produced, shared, evaluated, and legitimised within organisations. Building on insights from the knowledge management, organizational behaviour, and information systems literature, this study develops an integrated conceptual framework that links technological design features to employee behaviour and knowledge-sharing outcomes through embedded ethical risks.

Adopting an integrative and theory-building methodology, this research synthesises interdisciplinary literature to identify recurring mechanisms through which digital KM infrastructures influence organisational dynamics. The analysis highlights three interrelated dimensions of ethical risk: (1) surveillance and visibility, (2) control and standardisation, and (3) power asymmetries and transparency. These dimensions are conceptualised as embedded governance mechanisms that regulate not only the flow of knowledge but also the behavioural and epistemic conditions under which employees engage in knowledge-sharing practices.

The findings show that increased visibility and traceability, while enhancing coordination and accountability, may reduce psychological safety when associated with monitoring and evaluation, leading to self-censorship and selective participation. Similarly, standardisation and codification processes tend to privilege explicit and measurable knowledge, contributing to epistemic narrowing and limiting organisational knowledge diversity. Finally, algorithmic opacity and information asymmetries undermine perceptions of procedural fairness and organisational trust, weakening employees' willingness to contribute. These dimensions interact to produce reinforcing effects that encourage defensive behaviours, such as knowledge hiding and risk-averse participation, ultimately constraining organisational learning.

By linking these technological features to micro-level mechanisms, such as trust, intrinsic motivation, perceived fairness, and psychological safety, this study demonstrates that ethical risks are not external constraints but endogenous components of KM systems that shape knowledge-sharing outcomes. This study advances four theoretical propositions that explain how these mechanisms operate individually and in combination.

This study contributes to KM theory by integrating behavioural and ethical perspectives into the analysis of digital infrastructures, reframing KM systems as instruments of knowledge governance. It also provides practical implications, suggesting that organisations should prioritise transparency, participatory design, and psychologically safe environments to foster more inclusive and sustainable knowledge-sharing practices.

Keywords: Knowledge management; Knowledge sharing; Digital technologies, Organizational behaviour; Ethical issues; KM; Employee behaviour

Paper type: Academic Research Paper

1 Introduction

In contemporary organisations, digital technologies increasingly function as core infrastructure through which knowledge is produced, circulated, evaluated, and institutionalised (Karasti et al., 2016). Knowledge management (KM) systems, collaborative platforms, algorithmic decision-support tools, and AI-based repositories no longer merely support knowledge-related activities; they actively contribute to shaping how knowledge is rendered visible, legitimate, and actionable within organizational contexts (Dipasquale et al. 2024).

Despite this growing centrality, much of the KM literature has traditionally conceptualised technology as an enabling and largely neutral instrument, emphasising its potential to enhance efficiency, innovation, and organizational learning (Yan et al., 2025). While this instrumental perspective has been analytically productive, it tends to underplay how the integration of digital technologies into everyday knowledge practices reconfigures the conditions under which knowledge is produced and used within organizations.

Research on organizational behaviour and information systems suggests that technologies do not operate in isolation from social contexts but interact with norms, expectations, and behavioural patterns within organisations (Mueller et al., 2016). From this perspective, digital knowledge infrastructures shape not only the circulation of knowledge but also the conditions under which it is shared, evaluated, and recognised within organizational settings. However, despite the growing recognition of these dynamics across adjacent fields, their ethical and behavioural implications remain weakly integrated into mainstream KM theorising. Ethical concerns are often treated as secondary or normative add-ons rather than analytically relevant mechanisms shaping knowledge-related behaviours (Hong et al., 2023). As a result, the link between digital KM technologies, employee behaviour, and organizational learning outcomes remains insufficiently investigated. Against this backdrop, this study aims to critically examine technology-enabled KM systems as potential instruments of knowledge governance (Buuren, 2009), focusing on the ethical risks embedded in their design and use. Rather than framing ethics as an external constraint, this study conceptualises ethical risks as organizational features that operate endogenously, shaping employee behaviour, trust, and willingness to engage in knowledge sharing (Wang, 2016; Lin, 2007).

The primary objective is to develop an integrated conceptual framework that explains how ethical risks embedded in digital KM infrastructures influence key behavioural processes (Khanal, 2025) —such as perceived fairness, psychological safety, intrinsic motivation, and organizational trust — and, in turn, affect the quality and sustainability of knowledge sharing within organisations (Durst and Zieba, 2019). In doing so, this study advances a behaviourally grounded and organizationally informed understanding of technology-enabled KM, contributing to ongoing debates on the governance implications of digital knowledge infrastructures.

2 Theoretical Framing

In contemporary organisations, digital knowledge management systems operate as socio-technical infrastructures that shape how knowledge is shared, evaluated, and made actionable (Thomas, 2024), embedding behavioural and ethical implications within everyday work practices. These dynamics become particularly salient in knowledge-intensive settings, where knowledge sharing relies heavily on relational and psychological conditions such as trust, intrinsic motivation, and psychological safety (Lin, 2007; Wang, 2016). Prior research in organizational behaviour shows that when work practices become more visible, standardised, or subject to automated evaluation, employees tend to adjust their behaviours in ways that affect openness, voice, and willingness to contribute (Sherf et al., 2020).

Applied to technology-enabled KM systems, this suggests that increased traceability, codification, and algorithmic mediation may simultaneously support coordination while introducing new tensions into the social foundations of collective learning (Sánchez-Vera, 2025). On the one hand, digital infrastructures can enhance access to knowledge and organizational memory (Borgman et al., 2019); on the other hand, they may reshape how individuals perceive risks associated with sharing knowledge, how they interpret evaluation processes, and how they position themselves within organizational knowledge dynamics (Gibson et al., 2022).

This perspective highlights the need to move beyond a purely functional view of KM technologies and consider their role as organizational mechanisms that structure both behavioural responses and epistemic processes. In particular, technological features such as visibility and traceability, codification and standardisation, and algorithmic evaluation can be understood as distinct but interrelated dimensions through which digital infrastructures shape the production, interpretation, and valuation of knowledge within organisations (Alavi and Leidner, 2001; Turulja and Bajgoric, 2018).

3 Methodology

This study adopts an integrative conceptual approach aimed at theory development through a systematic synthesis of the existing literature. Rather than conducting a purely systematic review or narrative overview, this study follows an integrative logic that combines insights from multiple research streams to build a theoretically grounded framework (Bonte et al., 2024) on the ethical and behavioural implications of technology-enabled KM systems.

The literature selection process was guided by the objective of capturing interdisciplinary contributions relevant to the intersection of knowledge management, organizational behaviour, human resource management, and information systems. Seminal and recent studies were identified through targeted searches in major academic databases (i.e., Scopus and Web of Science), complemented by backward and forward citation tracking to ensure coverage of both foundational works (Hirt et al., 2021) and emerging debates on digital technologies and algorithmic governance in KM contexts (Turulja and Bajgoric, 2018).

Moreover, the selection of the proper works followed a relevance-based logic, focusing on contributions that explicitly addressed the role of digital technologies in shaping knowledge processes, employee behaviour, and organizational dynamics. Particular attention was paid to studies examining visibility, control mechanisms, algorithmic systems, and their implications for trust, motivation, and knowledge sharing.

The analytical process consisted of iterative and abductive readings of the selected literature to identify recurring themes, conceptual tensions, and underlying mechanisms. Through this process, the study identified three interrelated dimensions, surveillance and visibility, control and standardization, and power asymmetries and opacity, as key ethical risk areas embedded in digital KM infrastructures.

These dimensions were not derived inductively from a single stream of literature but emerged from the cross-fertilisation of insights across disciplines. Thus, framework development follows a theory-building logic, in which concepts and mechanisms are progressively integrated and refined to explain how technological features influence micro-level behavioural processes and, ultimately, knowledge-sharing outcomes.

By adopting this integrative approach, this study moves beyond the fragmented analyses of KM technologies and provides a coherent conceptual model that connects technological design, ethical risk, and organizational behaviour.

4 Findings

The literature review identifies three interrelated dimensions of ethical risk in technology-enabled KM systems. Rather than functioning solely as infrastructures for knowledge storage and exchange, these systems operate as sociotechnical arrangements that shape the behavioural, epistemic, and power conditions of knowledge sharing. The analysis highlights three key dimensions-surveillance and visibility, control and standardisation, and power asymmetries and opacity-which jointly configure how knowledge is produced, evaluated, and circulated within organisations. These dimensions do not act in isolation; instead, they function as embedded governance mechanisms that influence both the content of knowledge shared and the conditions under which employees are willing to contribute.

Surveillance and Visibility: Regulating Knowledge Sharing Behavior

Digital KM systems increasingly render knowledge-related activities visible, traceable, and auditable, thus transforming knowledge sharing into an observable organizational practice. This visibility enhances coordination and accountability by making expertise, contributions, and interactions more transparent (Varis et al., 2023). However, research consistently points to a fundamental visibility–behaviour paradox in this context.

When visibility is perceived as linked to monitoring or performance evaluation, employees engage in self-regulatory behaviours to minimise personal risk (Sherf et al., 2020). These include self-censorship, selective participation, and a preference for sharing “safe” and noncontroversial knowledge. Thus, knowledge sharing shifts from a collaborative activity to strategically managed behaviour shaped by concerns over reputation, evaluation, and potential misuse.

This dynamic undermines psychological safety and reduces the willingness to engage in voice behaviour, with potential consequences for knowledge diversity and organizational learning. In this sense, visibility operates as a mechanism of behavioural regulation, capable of both enabling coordination and constraining openness, depending on how it is designed and governed

Control and Standardization: Regulating What Counts as Knowledge

The second dimension concerns the increasing reliance of KM systems on codification and standardisation, which privilege knowledge that is explicit, structured, and measurable (Yan et al., 2025). While these features support efficiency, comparability, and scalability, they introduce a distinct ethical risk related to the epistemic structuring of organizational knowledge.

The literature highlights how KM systems embed an architectural bias toward explicit knowledge, marginalising tacit, contextual, and experiential forms (Garrick and Chan, 2017). This results in a process of epistemic filtering, whereby only knowledge that can be codified within predefined templates becomes visible and legitimate in the organisation.

This dynamic extends beyond knowledge storage to shape employees’ behaviour. Individuals adapt their contributions to fit standardised formats, aligning their thinking with system requirements rather than the intrinsic complexity of the phenomena they seek to describe. Over time, this leads to epistemic narrowing, reducing the diversity of perspectives and constraining exploratory and reflective forms of sense-making (Introne et al., 2020).

Thus, standardisation functions as a mechanism of epistemic regulation, defining what knowledge "counts" organizationally and subtly redirecting knowledge production toward measurable and system-compatible forms of knowledge.

Power Asymmetries and Opacity: Regulating Knowledge Valuation and Authority

The third dimension concerns the emergence of power asymmetries and opacity associated with algorithmic and AI-based KM systems. These technologies redistribute decision-making authority toward system designers, managers, and algorithmic processes, creating what has been described as a form of algorithmic governance or "algocracy" (Vladimir, 2025).

A central feature of this transformation is structural opacity. Algorithmic systems often operate as black boxes, making it difficult, if not impossible, for employees to understand how their knowledge is evaluated, ranked or utilised (Ajunwa, 2020). This opacity generates significant information asymmetries, whereby organizational actors controlling the system possess substantially more knowledge than those who are subject to it (Bergh et al., 2019).

These asymmetries produce both behavioural and relational consequences for the parties involved. From a behavioural perspective, opacity limits employees' ability to strategically engage with KM systems because the criteria governing evaluation remain unclear. From a relational perspective, it undermines perceptions of procedural fairness and organizational trust because employees cannot verify the impartiality or legitimacy of algorithmic decisions (Lee, 2018).

Moreover, literature highlights the risk of epistemic injustice, whereby algorithmic systems systematically privilege certain forms of knowledge while marginalising others, often reproducing existing organizational biases. In this sense, power asymmetries and opacity function as mechanisms of authority concentration, shaping not only how knowledge is evaluated but also who is recognised as a legitimate knowledge contributor.

Integrated Mechanisms: From Ethical Risk to Knowledge-Sharing Outcomes

The three dimensions identified in this study, surveillance and visibility, control and standardization, and power asymmetries and opacity, configure a multi-layered system through which technology-enabled KM shapes knowledge-sharing processes. These dimensions operate jointly to influence not only the extent to which employees are willing to share knowledge but also the forms that such knowledge can take and the criteria through which it is evaluated and recognised within the organisation.

This integrated configuration works through a set of micro-level behavioral mechanisms, most notably psychological safety, perceived fairness, organizational trust, and intrinsic motivation, which seem to mediate the relationship between KM infrastructure and knowledge-sharing outcomes (Yin et al., 2020). In this sense, the ethical risks embedded in digital KM systems are not external constraints but endogenous factors that shape how individuals interpret and engage with knowledge-sharing practices.

When these mechanisms are undermined, for instance, through excessive visibility linked to evaluation, rigid standardisation that constrains expression, or opaque systems that obscure decision-making processes, employees are more likely to adopt defensive behaviours such as self-censorship, knowledge hiding, and risk-averse participation. Although these responses are individually rational, they may collectively reduce knowledge diversity, limit exploratory learning, and weaken the overall effectiveness of knowledge management processes.

Conversely, when such ethical risks are mitigated through transparent governance arrangements, participatory system design, and organizational environments that support trust and psychological safety, KM systems can sustain more open and generative knowledge exchange. Under these conditions, digital infrastructure contributes to the development of more resilient and inclusive knowledge ecosystems.

5 Discussion and conclusion

The findings of this study support a theoretical reframing of technology-enabled knowledge management (KM) systems, shifting their interpretation from neutral infrastructures that facilitate knowledge processes to organizational arrangements that actively shape the behavioural and epistemic conditions under which knowledge is produced, shared and evaluated (Cristofaro and Bañón-Gomis, 2026). While the KM literature has traditionally emphasised the enabling role of digital technologies in supporting coordination and organizational learning (Alavi & Leidner, 2001), this study shows that these systems simultaneously embed normative assumptions, control mechanisms, and power configurations that influence how individuals engage with knowledge practices.

This perspective suggests that KM systems seem like forms of embedded governance, structuring not only the circulation of knowledge but also the conditions of participation, recognition, and evaluation of knowledge.

In this sense, ethical risks are not external to KM systems but are intrinsic features of their design and operation. Building on this reframing, this study conceptualises these dynamics as a set of interrelated mechanisms through which ethical risks shape knowledge-sharing processes.

The findings indicate that the effects of KM systems on knowledge sharing cannot be understood by isolating single technological features but must be interpreted as the outcome of interacting mechanisms that simultaneously shape behaviour, knowledge boundaries, and evaluative processes.

Based on these findings, this study advances a set of theoretical propositions that articulate how the identified dimensions translate into knowledge-sharing outcomes.

Proposition 1.

When visibility and traceability in KM systems are associated with monitoring and evaluation, they tend to reduce psychological safety and promote self-regulatory behaviours, such as self-censorship and selective knowledge sharing. This finding suggests that the effectiveness of visibility depends on how it is interpreted by organizational members: when framed as control rather than coordination, it shifts knowledge sharing from an intrinsically motivated activity to risk-managed behaviour.

Proposition 2.

The codification and standardisation embedded in KM systems tend to privilege explicit and measurable knowledge, contributing to the processes of epistemic narrowing. As a result, employees are likely to align their contributions with system-compatible formats, reducing the diversity and exploratory nature of organizational knowledge. This implies that KM systems do not merely capture knowledge but actively shape its form, favouring standardisation over interpretive richness and thereby constraining sensemaking processes.

Proposition 3.

The opacity and information asymmetries associated with algorithmic KM systems undermine perceived procedural fairness and organizational trust, thereby weakening employees' willingness to engage in knowledge-sharing practices. This highlights that trust in KM systems is not only a function of outcomes but also of the perceived legitimacy and transparency of the processes through which knowledge is evaluated and valued.

Proposition 4.

The interaction of surveillance, standardisation, and opacity produces reinforcing effects that amplify defensive behaviours, knowledge hiding, and risk aversion, ultimately constraining organizational learning. This suggests that ethical risks operate systemically, with combined effects that exceed the impact of individual dimensions and generate self-reinforcing dynamics.

Thus, these propositions point to a paradoxical dynamic: KM systems designed to enhance knowledge circulation may, under certain governance configurations, constrain openness, reduce the diversity of contributions, and limit learning processes. Taken together, these insights suggest that the effectiveness of technology-enabled KM systems depends not only on their technical capabilities but also on the configuration of ethical and governance conditions that shape their use.

From a practical standpoint, these findings indicate that organisations should carefully consider how visibility, codification, and algorithmic evaluation are implemented in KM systems. Overly rigid or opaque configurations may unintentionally discourage knowledge sharing, whereas approaches that emphasise transparency, participatory governance, and psychological safety may foster sustainable and inclusive knowledge practices.

5.1 Theoretical Implication

This study contributes to the advancement of knowledge management theory by introducing a more integrated understanding of the relationship between technology, behaviour, and epistemic structures. First, it extends the literature of KM, highlighting the role of digital technologies in structuring organizational processes beyond efficiency and performance, highlighting their capacity to regulate behaviour, define knowledge boundaries and redistribute authority.

Second, this study strengthens the connection between KM and organizational behaviour by explicitly theorising the micro-level mechanisms through which technological systems influence knowledge sharing. While prior research has acknowledged the importance of factors such as trust, motivation, and psychological safety (Wang, 2016), these elements have often been treated as contextual variables rather than outcomes of technological design. By demonstrating how visibility, standardisation, and opacity shape these mechanisms, this study repositions them as endogenous to KM systems rather than as external moderators.

Third, this study introduces an ethical–epistemic perspective to KM research, emphasising that technologies not only influence how knowledge is shared but also what knowledge is recognised and valued. This perspective contributes to emerging debates on the "dark side" of digital technologies (Aldhafferi, 2018) by demonstrating

that KM systems can simultaneously enable and constrain knowledge processes, depending on how they structure visibility, codification, and control.

In doing so, this study positions the issues of epistemic diversity, fairness, and inclusion as central theoretical concerns within KM research.

Finally, the framework developed in this study offers a basis for future research. Identifying ethical risks as mechanisms that mediate the relationship between KM technology and knowledge sharing outcomes opens new avenues for empirical investigation.

Future research may explore how different configurations of KM systems shape employee behaviour across organizational contexts, how governance arrangements influence these dynamics, and how organisations can balance efficiency with inclusiveness and epistemic diversity.

Future research could investigate sector-specific variations in KM governance, the role of leadership in mediating ethical risks, and the long-term implications of algorithmic KM systems for organizational learning and innovation.

6 Ethics Declaration

No ethical approval was required for this study. The paper is based on a conceptual and theory-building approach that relies exclusively on the analysis and integration of existing literature. It does not involve human participants, personal data, or any form of empirical data collection.

7 AI Declaration

The authors declare that generative artificial intelligence (AI) tools, specifically ChatGPT, were used solely to improve the clarity and quality of the English language in the manuscript. The authors' native language is Italian, and AI support was limited to linguistic refinement. All ideas, arguments, and intellectual contributions remain entirely those of the authors.

References

Ajunwa, I. (2020) "The "black box" at work" *Big Data & Society*, Vol. 7 No. 2, p. 2053951720966181.

Aldhafferi, N. (2018) "The dark side of digital technologies" In *2018 21st Saudi Computer Society National Computer Conference (NCC)* (pp. 1-6). IEEE.

Alavi, M., & Leidner, D. E. (2001) "Knowledge management and knowledge management systems: Conceptual foundations and research issues", *MIS quarterly*, pp. 107-136.

Bergh, D. D., Ketchen Jr, D. J., Orlandi, I., Heugens, P. P., & Boyd, B. K. (2019) "Information asymmetry in management research: Past accomplishments and future opportunities" *Journal of management*, Vol. 45 No. 1, pp. 122-158.

Bonte, P., Calbimonte, J. P., de Leng, D., Dell'Aglio, D., Della Valle, E., Eiter, T., & Ziffer, G. (2024) "Grounding stream reasoning research" *Transactions on Graph Data and Knowledge (TGDK)*, Vol. 2 No. 1, pp. 1-47.

Borgman, C. L., Scharnhorst, A., & Golshan, M. S. (2019) "Digital data archives as knowledge infrastructures: Mediating data sharing and reuse" *Journal of the Association for Information Science and Technology*, Vol. 70 No. 8, pp. 888-904.

Buuren, A. V. (2009) "Knowledge for governance, governance of knowledge: Inclusive knowledge management in collaborative governance processes", *International public management journal*, Vol. 12, No. 2, pp. 208-235.

Cristofaro, M., & Bañón-Gomis, A. J. (2026) "Dancing with the algorithm: a framework to navigate knowledge and autonomy in AI-assisted managerial decisions" *Journal of Knowledge Management*, Vol. 30 No. 11, pp. 1-31.

Dipasquale, L., Ammendola, J., Montoni, L., Ferrari, E. P., & Zambelli, M. (2024) "Harnessing vernacular knowledge for contemporary sustainable design through a collaborative digital platform", *Heritage*, Vol. *7* No. 9, p. 5251.

Durst, S., & Zieba, M. (2019) "Mapping knowledge risks: towards a better understanding of knowledge management", *Knowledge Management Research & Practice*, Vol. 17, No. 1, pp. 1-13.

Garrick, J., & Chan, A. (2017) "Knowledge management and professional experience: the uneasy dynamics between tacit knowledge and performativity in organizations" *Journal of knowledge management*, Vol. *21* No. 4, pp. 872-884.

Gibson, C. B., Dunlop, P. D., Majchrzak, A., & Chia, T. (2022) "Sustaining effectiveness in global teams: The coevolution of knowledge management activities and technology affordances", *Organization Science*, Vol. 33, No. 3, pp. 1018-1048.

Hirt, J., Nordhausen, T., Appenzeller-Herzog, C., & Ewald, H. (2021) "Using citation tracking for systematic literature searching-study protocol for a scoping review of methodological studies and a Delphi study" *F1000Research*, Vol. 9, p. 1386.

Hong, Y., Kim, B. J., & Kim, M. J. (2023) "The knowledge-sharing implications of social responsibility of firms: the importance of ethical climate", *Behavioral Sciences*, Vol. 13 No. 7, p. 608.

Introne, J., Korsunska, A., Krsova, L., & Zhang, Z. (2020) "Mapping the narrative ecosystem of conspiracy theories in online anti-vaccination discussions" In *International Conference on social media and society*, pp. 184-192.

Karasti, H., Millerand, F., Hine, C. M., & Bowker, G. C. (2016) "Knowledge infrastructures: part I", *Science & Technology Studies*, Vol. 29 No. 1, pp. 2-12.

Khanal, K. (2025). Ethics and Governance in Knowledge Management. In *Knowledge Sharing and Fostering Collaborative Business Culture* (pp. 349-362). IGI Global Scientific Publishing.

Lee, M. K. (2018) "Understanding perception of algorithmic decisions: Fairness, trust, and emotion in response to algorithmic management" *Big data & society*, Vol. 5 No. 1, p. 2053951718756684.

Lin, C. P. (2007) "To share or not to share: Modeling tacit knowledge sharing, its mediators and antecedents", *Journal of business ethics*, Vol. 70 No. 4, pp. 411-428.

Mueller, B., Renken, U., & van Den Heuvel, G. (2016) "Get your act together: An alternative approach to understanding the impact of technology on individual and organizational behavior" *ACM SIGMIS Database: The DATABASE for Advances in Information Systems*, Vol. 47 No. 4, pp. 67-83.

Sánchez-Vera, F. (2025) Critical algorithmic mediation: Rethinking cultural transmission and education in the age of artificial intelligence" *Societies*, Vol. 15 No. 7, p. 198.

Thomas, A. (2024) "Digitally transforming the organization through knowledge management: a socio-technical system (STS) perspective" *European Journal of Innovation Management*, Vol. 27 No. 9, pp. 437-460.

Turulja, L., and Bajgoric, N. (2018) "Information technology, knowledge management and human resource management: Investigating mutual interactions towards better organizational performance", *VINE Journal of Information and Knowledge Management Systems*, Vol. 48 No. 2, pp. 255-276.

Vladimir, N. (2025) "The artificial intelligence influence on structure of power: Long-Term transformation" *Legal Issues in the digital Age*, Vol. 2, pp. 183-212.

Wang, W. T. (2016) "Examining the influence of the social cognitive factors and relative autonomous motivations on employees' knowledge sharing behaviors", *Decision Sciences*, Vol. 47 No. 3, pp. 404-436.

Yan, M. R., Hong, L. Y., & Hu, R. R. (2025) "Benchmarking knowledge management practices and digital empowerment for business excellence in Asia", *Knowledge Management Research & Practice*, pp. 1-15.

Yin, J., Ma, Z., Yu, H., Jia, M., & Liao, G. (2020) "Transformational leadership and employee knowledge sharing: explore the mediating roles of psychological safety and team efficacy" *Journal of Knowledge Management*, pp. 24 No. 2, pp. 150-171.

An Index-Based Approach to SME Sustainability Learning and Governance

Fabrizio Flavio Baldassarre[1], Claudio Rosito[1], Valerio Buono[1], Gennaro Durante[2], Lucrezia Maria De Cosmo[1]
[1]University of Bari Aldo Moro, Bari, Italy
[2]Forethinking S.r.l. SB Bari, Italy

Abstract

Small and medium-sized enterprises (SMEs) are increasingly exposed to sustainability expectations arising from value chains, finance, and reporting infrastructures. Yet their responses often remain fragmented. Existing explanations usually stress resource scarcity, but fragmentation is also a knowledge problem: firms may recognise the relevance of sustainability while lacking the cognitive frames, governance routines, and prioritisation devices needed to turn diffuse concerns into coordinated action. This paper examines whether a structured self-assessment architecture can be interpreted not only as a measurement instrument but also as a developmental knowledge device. The study is situated within the broader National FSINDEX Observatory and reports evidence from a consolidated analytical subset of 15 Italian SMEs assessed between 20 January and 17 February 2026. Empirically, the paper relies on a document-based comparative analysis of the transversal FSINDEX report already produced for the participating firms. FSINDEX is a guided self-assessment tool organised around 50 ordinal items across five areas: Vision and awareness, Brand and corporate culture, Market drivers, Circular economy and social issues, and Governance. The analysis combines overall scores, area-level distributions, cluster positioning, and the recurring quick wins suggested in the consolidated report. The sample shows an average maturity score of 166.3 out of 250, with a range from 138 to 190. The strongest areas are Brand and corporate culture and Circular economy and social issues, while Market drivers and Governance remain the weakest. The pattern suggests that sustainability is more advanced as awareness and operational sensitivity than as a formal model of governance and market positioning. Recommended quick wins converge around governance and compliance, reporting and key performance indicators, people and welfare, and supply-chain formalisation. This concentration indicates that the main bottleneck is not simple awareness, nor primarily technology adoption, but the organisational work required to codify dispersed sustainability knowledge, assign responsibility, stabilise metrics, and sequence action. The paper contributes to knowledge management and sustainability pedagogy research by showing how a self-assessment observatory can function as a learning infrastructure for SMEs. It also refines the debate on sustainability maturity tools by arguing that their relevance lies less in classificatory precision alone than in their capacity to make strategic priorities discussable, comparable, and actionable.

Keywords: Sustainability pedagogy; knowledge management; SMEs; organisational learning; strategic prioritisation.

Paper type: Academic Research Paper

1 Introduction

Small and medium-sized enterprises are increasingly drawn into sustainability transition dynamics through customer requirements, supply-chain expectations, financial relationships, and evolving reporting frameworks. Recent European developments, including the Corporate Sustainability Reporting Directive and the subsequent encouragement of voluntary SME-oriented reporting standards, confirm that the sustainability agenda now affects firms far beyond the population directly subject to formal reporting mandates. For SMEs, however, response patterns remain uneven. Prior research has linked this unevenness to familiar structural constraints such as scarce managerial time, limited financial slack, and weak formalisation (Revell et al., 2010; Martins et al., 2022; European Commission, 2025). These constraints matter, but they do not fully explain why organisations that already recognise sustainability as strategically relevant still struggle to convert that recognition into coherent priorities.

A second issue is interpretive rather than purely resource-based. Sustainability confronts decision makers with ambiguous signals, heterogeneous standards, and the possibility of symbolic or performative responses that improve external appearance more than organisational substance (Delmas and Burbano, 2011; Lyon and Montgomery, 2015; Bernini and La Rosa, 2024; Santos et al., 2024). In SMEs, where strategy is often concentrated around owner-managers and close operational circles, fragmented sustainability action may therefore be understood as a knowledge problem: how environmental, social, and governance concerns are noticed, interpreted, discussed, codified, and prioritised inside the firm.

This paper addresses that problem by linking sustainability pedagogy and knowledge management. The central proposition is that structured self-assessment can do more than classify maturity. Under certain conditions, it can operate as a learning device that externalises tacit judgements, creates a shared vocabulary, identifies gaps, and translates diffuse sustainability concern into prioritised organisational action. The paper examines this proposition through the National FSINDEX Observatory, a broader initiative currently involving 20 SMEs, from which the present paper analyses the first consolidated subset of 15 firms.

The study asks three questions. First, how can a structured sustainability pedagogy architecture support leadership competences for intelligent knowledge management in SMEs? Secondly, how are strategic priorities distributed across the five FSINDEX areas in the analysed sample? Thirdly, what does the recurrence of similar gaps and quick wins suggest about the role of an observatory setting as a shared learning infrastructure?

The paper makes two contributions. Theoretically, it connects responsible management competence development with knowledge-based explanations of how SMEs build the capacity to interpret and mobilise sustainability knowledge (Cohen and Levinthal, 1990; Grant, 1996; Laasch et al., 2023). Empirically, it provides evidence from a structured comparative analysis showing that many SMEs are not in a condition of sustainability absence, but in a transitional condition in which awareness has outpaced governance, metrics, and market translation. The argument is deliberately cautious. The paper does not claim longitudinal proof of competence change. Rather, it shows how the architecture of FSINDEX and the patterning of its outputs illuminate the organisational work through which sustainability becomes manageable.

2 Theoretical Background

2.1 Sustainability pedagogy and responsible management competencies

Education for sustainable development has increasingly been defined as a transformative process rather than the transmission of a stable body of content. UNESCO (2020) frames sustainability learning as a lifelong process integrating cognitive, social, emotional, and behavioural dimensions. In related work on transformative learning, sustainability-related change involves shifts in meaning perspectives, critical reflection, and the reconstruction of assumptions through experience and dialogue (Mezirow, 1991; Grund et al., 2024). Competence-oriented scholarship has translated this pedagogic insight into more operational terms. Sustainability competence frameworks emphasise systems thinking, anticipatory reasoning, normative judgement, strategic action, and interpersonal capabilities as the basis for addressing complex sustainability problems (Wiek et al., 2011; European Commission, Joint Research Centre, 2022; Brundiers et al., 2021).

Within management studies, this debate has evolved into a broader discussion on responsible management competences. Rather than reducing responsibility to compliance or ethical intention, this literature defines competence as a layered combination of intellectual, behavioural, and personal capabilities required to judge, negotiate, and act across ethical, social, and environmental tensions (Laasch et al., 2023). Such competences include future orientation, critical judgement, stakeholder sensitivity, reflexivity, and the capacity to act in situations marked by ambiguity and competing claims. This is particularly relevant for SMEs, where formal functional boundaries are often weak and strategic decisions depend on the interpretive capacity of a small number of actors.

A pedagogic approach to sustainability in SMEs must therefore be practical, contextual, and developmental. It is less likely to depend on formal classroom settings than on situated devices that provoke reflection, comparison, and action. The significance of assessment lies precisely here. If designed developmentally rather than punitively, assessment can become part of pedagogy by enabling managers to see relationships, surface blind spots, and sequence change. In this sense, sustainability pedagogy in SMEs is inseparable from the instruments through which knowledge is structured and discussed.

2.2 Knowledge management, organisational learning, and strategic prioritisation in SMEs

Knowledge management provides a complementary explanation for why sustainability transitions stall or consolidate. Organisational learning research has long shown that firms depend on processes of knowledge acquisition, distribution, interpretation, memory, and institutionalisation (Huber, 1991; Crossan et al., 1999; Argote and Miron-Spektor, 2011). The absorptive capacity tradition further clarifies that new knowledge rarely becomes useful without prior capabilities to recognise, assimilate, and apply it (Cohen and Levinthal, 1990; Zahra and George, 2002). For sustainability problems, this challenge is amplified because firms must connect external expectations, internal routines, and long-term implications under conditions of uncertainty and temporal tension (Bansal and DesJardine, 2014).

Knowledge-based views of the firm help explain why this is difficult. Organisational knowledge does not emerge automatically from information availability; it requires integration mechanisms, routines, and interpretive interaction (Grant, 1996; Nonaka, 1994; Nonaka and Takeuchi, 1995). In SMEs, the literature suggests both an advantage and a vulnerability. Close proximity between decision makers and operations can support rapid experimentation and the rapid circulation of practical knowledge, yet weak formalisation can leave learning highly person-dependent and difficult to stabilise (Bos-Brouwers, 2010; Johnson, 2017). Sustainability initiatives may thus appear in the form of isolated practices, motivated individuals, or customer-led adaptations without becoming a durable organisational capability.

This problem becomes visible in strategic prioritisation. Sustainability requires firms to decide which issues matter, which actions come first, and which routines deserve investment. That sequence is never neutral. It reflects how managers classify problems, perceive trade-offs between exploration and exploitation, and judge the relationship between internal readiness and external positioning (March, 1991; Feeney et al., 2023). When prioritisation fails, firms may oscillate between symbolic narratives and disconnected initiatives. Hence the growing relevance of self-assessment and maturity models aimed at SMEs. Existing contributions show that diagnostic tools can help structure reflection and identify developmental pathways, but they are often treated as classificatory instruments rather than as socio-cognitive infrastructures for learning (Vásquez et al., 2021; Murè et al., 2024; Cicchini et al., 2024).

The present paper develops this latter perspective. It interprets FSINDEX as a developmental knowledge device: an artefact that does not merely describe sustainability maturity but helps organise attention, comparability, and action. This interpretation is especially pertinent when the instrument is embedded in an observatory setting where repeated patterns across firms can support benchmarking and peer-oriented learning (Wenger, 1998; Osagie et al., 2018).

3 Research Design

3.1 Research setting and the National FSINDEX Observatory

The research is situated within the National FSINDEX Observatory, a broader initiative aimed at supporting SME sustainability learning through structured self-assessment, collective benchmarking, and targeted developmental guidance. While the wider observatory currently involves 20 SMEs, this paper reports the first consolidated evidence available for 15 firms whose FSINDEX materials were completed between 20 January and 17 February 2026. The analysed subset covers heterogeneous sectors, including footwear, agriculture, agro-food, agrochemicals, logistics, ICT, graphic industry, electrical materials, construction, and industrial manufacturing.

The observatory logic is important for interpretation. The participating firms are not treated as isolated cases alone but as members of a common learning architecture in which results can be read comparatively. The empirical material available for this paper does not include full longitudinal evidence, nor direct interview transcripts. What is available is the transversal analytical report consolidating the 15 firm-level outputs. Accordingly, the paper adopts a document-based comparative design and uses the broader observatory mainly as the conceptual and organisational setting within which the 15-firm subset is positioned.

3.2 FSINDEX as a developmental knowledge device

FSINDEX is a proprietary strategic self-assessment tool developed by Forethinking Srl SB. It is not designed as a certification scheme or an external ESG rating. Its declared purpose is diagnostic and developmental: to help SMEs structure sustainability-related knowledge and translate it into prioritised improvement trajectories. The instrument comprises 50 items, ten for each of five areas: Vision and awareness; Brand and corporate culture; Market drivers; Circular economy and social issues; and Governance. Each item is scored on an ordinal scale from 1 to 5, generating a maximum of 50 points per area and 250 overall.

The operational threshold used in the consolidated report is 39 points per area. This threshold is not treated as a strict pass-fail criterion but as an interpretive maturity benchmark. In practical terms, the instrument yields more than a score. It produces an area profile, highlights strengths and weaknesses, and associates them with suggested quick wins. In this respect, FSINDEX can be read as a device for codifying managerial perceptions, translating them into a common vocabulary, and exposing where organisational learning needs to occur first. Its analytical relevance therefore lies in the relationship between scores and the action logic that accompanies them.

Table 1: Research setting and FSINDEX design

Element	Specification
Research setting	National FSINDEX Observatory
Empirical subset	15 SMEs
Compilation period	20 January 2026 to 17 February 2026
Instrument structure	50 ordinal items, 10 per area, 1-5 scale
Areas	Vision and awareness; Brand and corporate culture; Market drivers; Circular economy and social issues; Governance
Maximum score	50 per area; 250 overall
Interpretive benchmark	39/50 per area
Outputs used in this paper	Overall scores, area scores, cluster positioning, recurring quick wins, directional scenario

3.3 Sample, materials, and analytical strategy

Methodologically, the paper adopts an embedded multiple-case comparative analysis based on the consolidated transversal report. The analysis uses three layers of evidence already integrated in that source: first, the numerical distribution of scores across the five areas; secondly, the cross-case positioning of firms by total score; thirdly, the pattern of quick wins associated with recurring developmental needs. The quick-win scenario is explicitly directional and non-predictive. It estimates plausible improvements in the areas directly activated by the recommended actions, with a cap at 50 points per area, and should therefore be read as a maturation trajectory rather than a forecast.

The analytical strategy is interpretive and comparative. Rather than testing causal relations statistically, it examines how the distribution of weaknesses and recommended actions reveals the organisational conditions under which sustainability knowledge becomes actionable. This approach is suitable for the present stage of the observatory because it allows cross-case synthesis without overstating what the available material can support.

4 Findings

4.1 Cross-sectional maturity profile of the 15 SMEs

The first result is that the sample does not depict an SME population starting from zero. The average total score is 166.3 out of 250, equivalent to 66.5% of the theoretical maximum. Scores range from 138 to 190, indicating substantial dispersion but not a polarity between highly mature and wholly inactive firms. Four firms fall into an advanced cluster with at least 180 points, six belong to a solid but non-homogeneous cluster between 165 and 179, two lie in a transitional range between 145 and 164, and three remain in a category requiring more substantial structuring below 145.

This distribution suggests an intermediate condition of sustainability maturity. Sustainability has entered managerial agendas across the sample, but the translation of this attention into stable systems remains uneven. The most developed cases are not necessarily those with complete formalisation in every area; rather, they are firms in which at least some domains of sustainability appear sufficiently structured to support an integrated profile. Conversely, the weakest cases still display areas of partial advancement, which indicates that even lower-scoring firms are engaged in sustainability-related practices, albeit in fragmented form.

The directional scenario reinforces this interpretation. If the suggested quick wins were implemented in a prudent manner, the average score would rise from 166.3 to 177.3, corresponding to an average gain of 11 points. The magnitude of this projected improvement is modest but meaningful. It implies that much of the near-term developmental potential lies not in radical transformation, but in converting existing awareness into more ordered and governable routines.

4.2 Recurring patterns across the five FSINDEX areas

The area distribution makes the central bottlenecks visible. The strongest average scores are recorded in Brand and corporate culture and Circular economy and social issues, both at 34.4 out of 50, followed closely by Vision and awareness at 34.2. Governance scores 33.0, while Market drivers is clearly the weakest area at 30.3. Only one firm exceeds the 39-point benchmark in Market drivers, whereas five do so in Governance.

Table 2: Area-level results in the analysed sample

Area	Mean	Firms above 39/50	Maximum	Minimum
Vision and awareness	34.2	3/15	42	26
Brand and corporate culture	34.4	4/15	43	27
Market drivers	30.3	1/15	41	21
Circular economy and social issues	34.4	3/15	41	29
Governance	33.0	5/15	43	20

Three patterns emerge from this distribution. First, sustainability is stronger as awareness and operational sensitivity than as a model of governance. The firms appear more prepared to recognise environmental and social impacts, discuss internal culture, and report selective practices than to anchor sustainability in formal roles, procedures, incentive systems, and control mechanisms. The transversal report explicitly contrasts relatively robust answers on risk awareness, energy, inclusion, safety, welfare, and environmental measurement with weaker answers on formal governance devices, structured stakeholder engagement, privacy, and compliance mechanisms.

Secondly, the internal symbolic and cultural side of sustainability appears more developed than its external market translation. The sample includes several firms whose culture-related scores signal openness to innovation, internal sensitivity, and employer attractiveness, yet these same firms often score poorly on eco-design, sustainable offer configuration, monetisation of sustainability, or customer recognisability of sustainability value. This gap is theoretically important. It indicates that sustainability knowledge may circulate internally as conviction or aspiration without yet becoming a stable value proposition.

Thirdly, the maturity profile is selective rather than systemic. The cross-sectional report notes the presence, in different firms, of life-cycle assessment, environmental product declarations, welfare initiatives, emissions monitoring, and early sustainability reporting efforts. These are not trivial signals. They suggest that sustainability has already generated technical and organisational experimentation. Yet the experiments often remain project-based or entrepreneur-led rather than translated into an integrated and durable management system. The sample therefore reveals a shift from marginality to partial institutionalisation, not full consolidation.

4.3 Quick wins as mechanisms of prioritisation and organisational learning

The quick wins provide the clearest indication that FSINDEX is functioning as more than a diagnostic scoreboard. Across the 15 firms, the suggested actions converge around a limited set of recurring levers: governance and compliance in 11 cases, reporting and key performance indicators in 10, people and welfare in 10, supply chain and procurement in 7, stakeholder engagement in 6, and product and innovation in 4.

Table 3: Recurring quick-win levers across the sample

Quick-win lever	Occurrences
Governance and compliance	11/15
Reporting and KPI	10/15
People and welfare	10/15
Supply chain and procurement	7/15
Stakeholder engagement	6/15
Product and innovation	4/15

The concentration is analytically revealing. Most recommended actions are not primarily technological. They concern formalisation, measurement, internal capability building, and accountability. In other words, the first developmental task for these SMEs is to convert sustainability from scattered initiative into organised knowledge. Quick wins do this by reducing complexity. They identify where responsibility should be assigned, where information should be stabilised in metrics, where people-related routines should be strengthened, and where value-chain visibility should be improved.

This helps explain the directional scenario. Governance is the area expected to improve most, followed by internal structuring related to Vision and awareness and supply-chain organisation. Market drivers improve less rapidly because commercial repositioning depends on prior internal ordering. A firm can rarely translate sustainability into a credible offer if its data, governance, and responsibilities remain unstable. The quick wins therefore perform a sequencing function. They do not simply recommend action; they establish the order in which action is likely to become effective.

5 Discussion

The findings support a reinterpretation of assessment in SME sustainability management. The main contribution of FSINDEX is not exhausted by the scores it generates. Its deeper value lies in its capacity to structure managerial attention. By transforming broad sustainability concerns into area-based profiles and associated quick wins, the instrument makes the transition from awareness to prioritisation visible and discussable. This is why the evidence is best understood through a knowledge management lens.

The first research question concerned leadership competences for intelligent knowledge management. The present cross-sectional material does not allow a direct measurement of competence change over time. Nonetheless, the patterns observed are consistent with the view that FSINDEX creates enabling conditions for competence development. The instrument requires managers to articulate sustainability judgements across multiple domains, confront asymmetries between internal culture and external positioning, and engage with actionable recommendations. These are not merely reporting acts. They involve interpretation, comparison, and decision sequencing, all of which are core components of intelligent knowledge management.

The second question asked how strategic priorities are distributed across the five areas. Here the evidence is direct. Priorities are concentrated less on radical product innovation than on the organisational foundations that make later strategic differentiation possible. The severe weakness of Market drivers, contrasted with comparatively higher scores in awareness and culture, shows that the central bottleneck is the conversion of sustainability knowledge into market-facing strategic architecture. From the perspective of absorptive capacity, the firms appear to have progressed further in recognising and valuing sustainability signals than in assimilating them into stable routines and exploiting them competitively (Cohen and Levinthal, 1990; Dzhengiz and Niesten, 2020).

This interpretation also aligns with organisational learning theory. The sample suggests movement from intuitive and project-based responses toward more institutionalised forms of learning, but with uneven progression across levels. Sustainability appears to have reached the stages of noticing, discussing, and selectively experimenting, yet often remains under-institutionalised in governance structures and performance systems (Crossan et al., 1999; Argote and Miron-Spektor, 2011). In Nonaka's terms, dispersed practical understandings are present, but the mechanisms required for durable conversion and integration remain incomplete (Nonaka, 1994; Nonaka and Takeuchi, 1995).

The third question concerned the observatory setting. Again, the present evidence is indirect but meaningful. The recurrence of similar quick wins across heterogeneous sectors suggests that the participating SMEs share a common developmental grammar. This is precisely what makes an observatory valuable. It does not merely collect cases; it reveals patterned bottlenecks that can be addressed through benchmarking, shared language, and peer-oriented learning. The concentration of recommended actions around governance, reporting, people, and supply chain implies that many SMEs can benefit from a relatively standardised developmental infrastructure before moving into more differentiated sustainability strategies. This resonates with research showing that competence development in sustainability-related roles is strongly supported by peer learning and learning climate rather than by formal instruction alone (Osagie et al., 2018).

The findings also have implications for the debate on symbolic compliance. Because FSINDEX is not designed as an external rating or legitimacy signal, it partly displaces the incentives associated with performative disclosure. Its logic is developmental rather than reputational. That does not eliminate the risk of self-report bias, but it does alter the purpose of the exercise. The central output is a prioritised improvement trajectory rather than a public score for external comparison. In this sense, the instrument may reduce some incentives associated with greenwashing by focusing managerial attention on internal gaps and next-step actions rather than external claims (Delmas and Burbano, 2011; Lyon and Montgomery, 2015).

Practically, the paper suggests that SMEs need lightweight but structured infrastructures for sustainability learning. The analysed firms do not primarily lack goodwill. They lack codified routines, comparable metrics, clear responsibilities, and pathways for converting sustainability into market and governance architecture. Instruments such as FSINDEX can therefore be valuable not because they simplify reality excessively, but because they render complexity governable at a scale compatible with SME conditions.

6 Conclusions

This paper has argued that FSINDEX is best understood not only as a sustainability maturity instrument but as a developmental knowledge device. Using a document-based comparative analysis of 15 SMEs within the broader National FSINDEX Observatory, the study showed that the main bottlenecks in the sample lie in Governance and especially Market drivers, while Vision and awareness, Brand and corporate culture, and Circular economy and social issues are comparatively more developed. The resulting picture is one of

intermediate maturity: sustainability is present, but often still fragmented, selectively formalised, and only partially translated into strategic priorities.

The paper contributes to the literature by linking sustainability pedagogy, responsible management competences, and organisational learning in an SME setting. Its main empirical insight is that self-assessment becomes strategically meaningful when it produces not just a diagnosis but a sequence of quick wins that codifies knowledge gaps and organises action. For practice, the implication is that support infrastructures for SMEs should prioritise governance, metrics, people-related routines, and supply-chain visibility as the foundations upon which stronger market-facing sustainability strategies can later be built.

The study has clear limitations. It relies on self-assessment material, which raises the possibility of self-report bias. It is cross-sectional rather than longitudinal and is based on a consolidated transversal report rather than raw firm-level protocols. The directional scenario attached to the quick wins is heuristic and should not be mistaken for a prediction. Future research should follow the broader observatory over time, incorporate direct qualitative evidence from participating firms, and test whether the learning dynamics suggested here translate into measurable changes in strategic behaviour, reporting quality, and market positioning.

7 Ethics Declaration

The paper reports a document-based comparative analysis of organisational self-assessment materials and aggregated cross-case evidence. No identifiable personal data are analysed in the manuscript, and no experimental procedures involving human participants are reported. No formal ethical clearance reference is reported for the present paper.

8 AI Declaration

OpenAI ChatGPT was used in the preparation of the present manuscript for language support, structural consolidation, and drafting assistance. All empirical inputs, source selection, interpretation of findings, reference checking, and final scholarly responsibility remain with the authors. No AI-generated material was included without human review and revision.

References

Durante, G. (2021) 'La transizione ecologica. Il percorso dell'impresa consapevole', Tecniche Nuove S.p.A.

Argote, L. and Miron-Spektor, E. (2011) 'Organizational learning: From experience to knowledge', Organization Science, 22(5), pp. 1123–1137.

Bansal, P. and DesJardine, M.R. (2014) 'Business sustainability: It is about time', Strategic Organization, 12(1), pp. 70–78.

Bernini, F. and La Rosa, F. (2024) 'Research in the greenwashing field: concepts, theories, and potential impacts on economic and social value', Journal of Management and Governance, 28, pp. 405–444.

Bos-Brouwers, H.E.J. (2010) 'Corporate sustainability and innovation in SMEs: Evidence of themes and activities in practice', Business Strategy and the Environment, 19(7), pp. 417–435.

Brundiers, K., Barth, M., Cebrián, G., Cohen, M., Diaz, L., Doucette-Remington, S., Dripps, W., Habron, G., Harré, N., Jarchow, M., Losch, K., Michel, J., Mochizuki, Y., Rieckmann, M., Parnell, R., Walker, P. and Zint, M. (2021) 'Key competencies in sustainability in higher education—toward an agreed-upon reference framework', Sustainability Science, 16(1), pp. 13–29.

Cicchini, D., Cotana, M.L., Filocamo, M.R. and Galeotti, R.M. (2024) 'Sustainability reporting practices in SMEs. A systematic study and future avenues', Piccola Impresa / Small Business, 1, pp. 104–134.

Cohen, W.M. and Levinthal, D.A. (1990) 'Absorptive capacity: A new perspective on learning and innovation', Administrative Science Quarterly, 35(1), pp. 128–152.

Crossan, M.M., Lane, H.W. and White, R.E. (1999) 'An organizational learning framework: From intuition to institution', Academy of Management Review, 24(3), pp. 522–537.

Delmas, M.A. and Burbano, V.C. (2011) 'The drivers of greenwashing', California Management Review, 54(1), pp. 64–87.

Dzhengiz, T. and Niesten, E. (2020) 'Competences for environmental sustainability: A systematic review on the impact of absorptive capacity and capabilities', Journal of Business Ethics, 162, pp. 881–906.

European Commission (2025) Commission presents voluntary sustainability reporting standard to ease burden on SMEs. Brussels: Directorate-General for Financial Stability, Financial Services and Capital Markets Union.

European Commission, Joint Research Centre (2022) GreenComp: The European Sustainability Competence Framework. Luxembourg: Publications Office of the European Union.

Feeney, M., Grohnert, T., Gijselaers, W. and Martens, P. (2023) 'Organizations, learning, and sustainability: A cross-disciplinary review and research agenda', Journal of Business Ethics, 184, pp. 217–235.

Grant, R.M. (1996) 'Toward a knowledge-based theory of the firm', Strategic Management Journal, 17(S2), pp. 109–122.

Grund, J., Singer-Brodowski, M. and Büssing, A.G. (2024) 'Emotions and transformative learning for sustainability: A systematic review', Sustainability Science, 19, pp. 307–324.

Huber, G.P. (1991) 'Organizational learning: The contributing processes and the literatures', Organization Science, 2(1), pp. 88–115.

Johnson, M.P. (2017) 'Knowledge acquisition and development in sustainability-oriented small and medium-sized enterprises: Exploring the practices, capabilities and cooperation', Journal of Cleaner Production, 142(Part 4), pp. 3769–3781.

Laasch, O., Moosmayer, D.C. and Antonacopoulou, E.P. (2023) 'The interdisciplinary responsible management competence framework: An integrative review of ethics, responsibility, and sustainability competences', Journal of Business Ethics, 187, pp. 733–757.

Lyon, T.P. and Montgomery, A.W. (2015) 'The means and end of greenwash', Organization & Environment, 28(2), pp. 223–249.

March, J.G. (1991) 'Exploration and exploitation in organizational learning', Organization Science, 2(1), pp. 71–87.

Martins, A.A., Madaleno, M. and Dias, M.F. (2022) 'Sustainability in SMEs: A systematic literature review and future research agenda', Sustainability, 14(11), 6493.

Mezirow, J. (1991) Transformative Dimensions of Adult Learning. San Francisco: Jossey-Bass.

Murè, P., Giorgio, S., Antonelli, V. and Crisafulli, A. (2024) '"ESG score" vs. "ESG rating": A conceptual model for the sustainability assessment and self-assessment of European SMEs', Frontiers in Environmental Economics, 3, 1452416.

Nonaka, I. (1994) 'A dynamic theory of organizational knowledge creation', Organization Science, 5(1), pp. 14–37.

Nonaka, I. and Takeuchi, H. (1995) The Knowledge-Creating Company. New York: Oxford University Press.

Osagie, E.R., Wesselink, R., Runhaar, P. and Mulder, M. (2018) 'Unraveling the competence development of corporate social responsibility leaders: The importance of peer learning, learning goal orientation, and learning climate', Journal of Business Ethics, 151, pp. 891–906.

Revell, A., Stokes, D. and Chen, H. (2010) 'Small businesses and the environment: Turning over a new leaf?', Business Strategy and the Environment, 19(5), pp. 273–288.

Santos, C., Coelho, A. and Marques, A. (2024) 'A systematic literature review on greenwashing and its relationship to stakeholders: State of art and future research agenda', Management Review Quarterly, 74, pp. 1397–1421.

UNESCO (2020) Education for Sustainable Development: A Roadmap. Paris: UNESCO.

Vásquez, J., Aguirre, S., Puertas, E., Bruno, G., Priarone, P.C. and Settineri, L. (2021) 'A sustainability maturity model for micro, small and medium-sized enterprises (MSMEs) based on a data analytics evaluation approach', Journal of Cleaner Production, 311, 127692.

Wenger, E. (1998) Communities of Practice: Learning, Meaning, and Identity. Cambridge: Cambridge University Press.

Wiek, A., Withycombe, L. and Redman, C.L. (2011) 'Key competencies in sustainability: A reference framework for academic program development', Sustainability Science, 6, pp. 203–218.

Zahra, S.A. and George, G. (2002) 'Absorptive capacity: A review, reconceptualization, and extension', Academy of Management Review, 27(2), pp. 185–203.

An Integrated Framework for Evaluating Telemedicine Systems in Regional Healthcare Contexts

Lorella Cannavacciuolo[1], Roberta De Cristofaro[1], Vincenzo Maione[2], Cristina Ponsiglione[1], Simonetta Primario[3]
[1] University of Naples Federico II, Naples, Italy
[2] LIUC, Cattaneo University, Castellanza, VA, Italy
[3] Pegaso Telematic University, Italy

Abstract

Telemedicine has progressively evolved from a complementary digital service into a structural component of contemporary care models. This transition has transformed telemedicine into a complex intervention whose effectiveness depends not only on technological performance, but also on the dynamic interactions among organizational processes, professional practices, patient engagement, infrastructural conditions, and governance arrangements. Although established multidimensional frameworks, such as Health Technology Assessment (HTA), provide robust methodological foundations for evaluating health technologies, their application at the regional ecosystem level still presents relevant challenges. In particular, traditional assessment approaches often struggle to capture the adaptive nature, contextual variability, stakeholder interdependencies, and systemic trade-offs that characterize telemedicine as a socio-technical system. Against this background, a persistent gap remains between the theoretical recognition of telemedicine complexity and the availability of operational tools capable of monitoring such complexity in an integrated and decision-oriented manner. This study addresses this gap by proposing the Telemedicine Ecosystem Monitoring and Evaluation Framework (TEMEF), an HTA-informed and context-sensitive model designed to support the evaluation of telemedicine systems in regional healthcare contexts. Rather than replacing existing evaluative approaches, TEMEF aims to complement them by translating multidimensional assessment principles into a structured, multi-stakeholder monitoring architecture. Using a Proof of Concept (PoC) methodology, the paper describes the transition from conceptual design to the development of a survey-based evaluation system tailored to four key stakeholder groups: patients and caregivers, physicians, healthcare managers, and technology providers. The framework integrates five main evaluative domains—clinical, organizational, technological, experiential, and economic—allowing system performance to be examined both at an aggregate level and through stakeholder-specific perspectives. To assess the internal coherence of the scoring logic and the interpretability of the analytical outputs, AI-generated synthetic stakeholder profiles were used to conduct a preliminary stress-test of the framework and its analytical dashboard interface. This procedure enabled the exploration of plausible configurations of stakeholder alignment and misalignment, highlighting how apparently positive aggregate performance may conceal critical tensions across domains or actor groups. The resulting model provides policy makers and healthcare decision-makers with an operational framework for interpreting telemedicine not as an isolated technological solution, but as a regional healthcare ecosystem. TEMEF supports evidence-informed governance by making visible systemic trade-offs, implementation bottlenecks, and areas requiring targeted improvement, thereby contributing to the sustainable integration of telemedicine into evolving care delivery models.

Keywords: Telemedicine, Healthcare Innovation, Multistakeholder Framework, Exploratory Study, Regional Healthcare Systems

Paper type: Academic Research Paper

1 Introduction

In the contemporary landscape of public health, digital transformation represents one of the primary strategic directions for investment and the development of healthcare systems (Rabbani et al., 2025). The implementation of increasingly sophisticated Information and Communication Technology infrastructures does not merely respond to a need for technological modernization; rather, it aims for a strategic objective: optimizing the efficiency and effectiveness of service delivery. Within this scenario, telemedicine has established itself as the solution of choice for mitigating the factor of distance, which has historically represented a critical barrier to equity in access to care. Due to its ability to overcome geographical boundaries and facilitate contact between physicians and patients, it has long been at the center of international policy agendas as a key instrument of e-health (Ricci et al., 2023) However, telemedicine has now moved beyond being a simple "alternative" to in-person visits to become a pillar of modern care models. It enables proactive care pathways and ensures genuine integration between the hospital environment and the patient's home, addressing the challenges posed by chronic diseases and the demand for proximity medicine (Li et al., 2025; Ricci et al., 2023;

Rogers et al., 2022). This transition, accelerated by the pandemic emergency and now crystallized in national digitalization plans, has highlighted that telemedicine cannot be reduced to a mere technological artifact. How, then, do we define the value of an ecosystem where digital infrastructures, professional competencies, patient needs, and governance rules interact in a dynamic and unpredictable manner? The answer lies in recognizing telemedicine as a complex socio-technical system: a living environment whose performance depends not only on software but on the quality of the relationships and processes it enables (Greenhalgh et al., 2017; Khodadad-Saryazdi, 2021). It is precisely this intrinsic complexity that makes its implementation (Ross et al., 2016) and evaluation an ongoing methodological challenge. While the literature has produced numerous models to analyze individual aspects of this phenomenon, such contributions often remain fragmented (Chang, 2015; Souza et al., 2021). The difficulty emerges particularly when the unit of analysis shifts from a single local application to an entire regional system. The primary limitation of traditional models is not a lack of rigor, but rather the difficulty in capturing the interdependencies between various actors and the nuances of implementation contexts. This critical issue is confirmed by recent reflections in the field of HTA: the work of Baghbanian et al. (2025) highlights that current HTA, despite having progressively expanded its scope beyond traditional domains, continues to encounter difficulties in capturing the dynamic interactions, contextual factors, and long-term impacts associated with health interventions. Based on this premise, the present study introduces the Telemedicine Ecosystem Monitoring and Evaluation Framework (TEMEF). This is not a proposal intended to replace existing frameworks, but rather an HTA-informed and context-sensitive approach designed to operationalize systemic complexity. The contribution of this work lies in translating theoretical constructs into a multi-stakeholder structure that recognizes and values the diverse perspectives of patients, physicians, managers, and providers. In this perspective, regional telemedicine is treated as a complex intervention and a multi-level socio-technical system. Its evaluation cannot be limited to the isolated measurement of clinical outcomes or technological performance; it must also include stakeholder interdependencies, implementation feasibility, adaptive capacity, contextual variability, and potential tensions between objectives of efficiency, quality, accessibility, and sustainability. In light of these premises, the paper pursues a dual objective. On one hand, it intends to contribute to the debate on the limitations of current evaluative approaches, showing how regional telemedicine requires models capable of explicating system complexity rather than merely summing distinct evaluative domains. On the other hand, it proposes an initial operational response to this need through a multi-stakeholder framework that translates dimensions of system complexity into empirically observable constructs. In this sense, the contribution of this work lies not in proposing yet another multidimensional frame for telemedicine, but in the attempt to bridge the gap between the theoretical recognition of complexity and its effective operationalization into evaluative tools applicable to regional contexts, providing a support tool for policy decision-makers to systematically evaluate system performance through the use of concrete indicators measurable over time.

2 Theoretical Background

The healthcare system is a fundamental pillar of socio-economic development, yet it faces unprecedented political, economic, and social pressures (Chiuchisan et al., 2014; Kelly & Young, 2017). Today's global environment—characterized by climate change, migration flows, and environmental degradation—challenges the resilience of traditional healthcare models, necessitating an accelerated pace of adaptation (Benis et al., 2021; Kim et al., 2014). Within this context, digital transformation emerges as a critical driver for reshaping organizational models toward greater integration and sustainability (Konopik, 2023; Reis et al., 2018). While digital platforms are now essential for orchestrating data exchange in systems that historically lacked integrated structures (Baby & Chandran, 2022; Essen et al., 2023; Gleiss et al., 2021), their implementation remains challenging due to the multidimensional and interconnected nature of the healthcare environment (Konopik, 2023). These challenges have led to the progressive adoption of Complexity Thinking in healthcare (Khan et al., 2018). Recognizing that healthcare is inherently complex (Plsek & Greenhalgh, 2001), it requires adaptive responses rather than traditional reductionist solutions (Ashby, 1991). Consequently, healthcare is increasingly described as a complex system (Begun JW et al., 2003), in which the connectivity between actors generates interdependencies, and the actions of a single participant can create 'ripple effects' throughout the entire structure. This complexity is further enhanced by the system's hierarchical organization across nested micro, meso, and macro levels (Khan et al., 2018). Addressing global issues such as limited access, operational inefficiencies, and rising costs (Yeganeh, 2019), therefore, requires innovative and scalable solutions—like telemedicine—capable of navigating these intricate interdependencies to improve the quality of care (Tabish & Nabil, 2015). In this context, Sood et al. (2007) summarize the broad concept of telemedicine in this definition "*Telemedicine being a subset of telehealth, uses communications networks for delivery of healthcare services*

and medical education from one geographical location to another, primarily to address challenges like uneven distribution and shortage of infrastructural and human resources." Although telemedicine has been part of the healthcare discourse for decades, its implementation expanded dramatically during the COVID-19 pandemic, acting as a mandatory catalyst for digital health adoption (Rabbani et al., 2025). Despite this widespread implementation, its evaluation remains a significant methodological challenge. Traditional evaluation frameworks, most notably HTA have provided a robust basis for assessing medical technologies through multidimensional lenses, such as clinical effectiveness and economic sustainability (Kidholm et al., 2012; Kristensen et al., 2017). However, as telemedicine shifts from a niche tool to a structural component of regional healthcare, traditional HTA models often encounter difficulties in capturing the full spectrum of its systemic impact. Current evaluative efforts indicate that telemedicine assessment has largely evolved along sectoral and partial trajectories. Some research streams prioritize clinical performance and health outcomes (Gonçalves-Bradley et al., 2018; Zhang et al., 2025), while others focus on information-system characteristics and technology-mediated interaction (Zhang et al., 2025). Still, other studies emphasize data quality, operational processes, and implementation requirements (Talwar et al. 2023). Although these specializations yield valuable insights, such contributions often remain fragmented and difficult to compare, especially when the unit of analysis transitions from a singular service or pilot initiative to a complex regional system. This critical issue is echoed in recent reflections within the HTA field. Specifically, Baghbanian et al. (2025) point out that while HTA has expanded beyond its original clinical and economic domains, it still struggles to account for the dynamic interactions, contextual factors, and long-term sustainability inherent in complex interventions. In such systems, evaluation cannot be a static or reductionist process. Instead, it requires a "complexity-informed" approach capable of monitoring geographic variability, stakeholder engagement, and implementation feasibility (Baghbanian et al., 2025). For a regional telemedicine ecosystem to be truly performant, the evaluation must embrace a multi-stakeholder perspective, where the divergent views of patients, physicians, and managers are treated as essential diagnostic signals of systemic trade-offs and misalignments.

3 Methods

The innovative nature and systemic complexity of TEMEF necessitated a methodological approach structured into two macro-phases: conceptual design and functional validation, establishing this study as a PoC. As shown in Figure 1, the methodological process is divided into five key steps: literature review; definition of the framework's conceptual architecture; development of a multi-stakeholder survey; generation of "twin profiles" for internal pre-testing; and the development of a preliminary scoring logic and dashboard visualization.

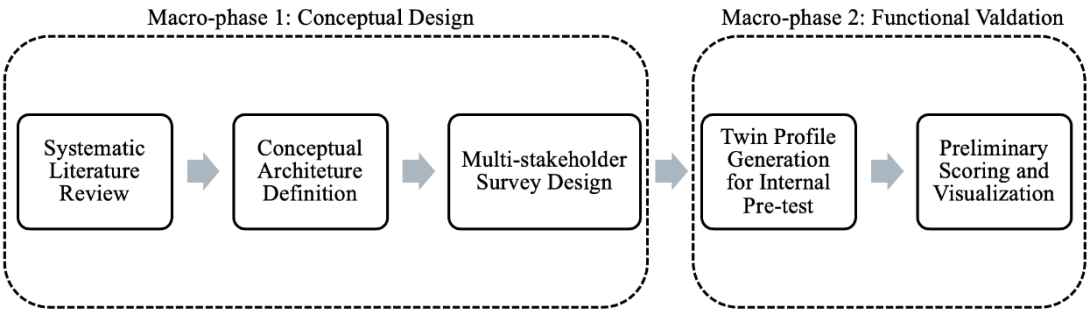


Figure 1: Methodological structure of the TEMEF PoC.

The framework stems from a systematic literature review conducted in Scopus using a query built around the terms "telemedicine" and "platform", selecting peer-reviewed journal articles written in English and published between 2020 and 2025 in subject areas consistent with the focus of the study, including engineering, information systems, management, economics, and decision sciences. This combination was chosen to prioritize contributions that analyze telemedicine through the lens of platform-based structures, which are increasingly recognized as essential for orchestrating data exchange and multi-stakeholder integration within complex healthcare environments. By focusing on "platforms", the review filtered for studies aligned with the socio-technical and systemic nature of the TEMEF, moving beyond traditional, fragmented views of e-health tools. This initial search returned 77 contributions. After title and abstract screening, 48 articles were selected as relevant. The excluded articles were not directly focused on telemedicine platforms, lacked an empirical or methodological basis, or were duplicates or partially overlapping studies. The included contributions were analyzed by extracting information on study context, technological focus, methodological approach, main

findings, and enabling or critical factors. These contributions helped us identify the main assessment dimensions considered in the literature, as well as the conceptual and evaluative models most frequently adopted, and confirmed a significant methodological gap: although theory widely recognizes the multidimensional nature of healthcare value, operational tools remain fragmented. Available evidence, while relevant, often appears difficult to compare and is rarely translatable into integrated decision-support systems—a critical issue that emerges prominently when the object of evaluation is a complex system within a regional healthcare context. To overcome this fragmentation, TEMEF integrates five fundamental domains—clinical, organizational, technological, experiential, and economic—into a unified architecture. The framework's development combined empirical literature with the analysis of established models (such as HTA, Donabedian's paradigm, the Chronic Care Model, NASSS, and UTAUT). While these provide solid foundations for interpreting telemedicine as a socio-technical system, they exhibit a recurring limitation: the difficulty of operationalizing complex theoretical constructs into empirically observable indicators. Consequently, the framework remains consistent with established evaluative traditions, including HTA, but distinguishes itself by translating these dimensions into a multi-stakeholder structure tailored for regional telemedicine systems. To translate these theoretical constructs into observable variables, a differentiated survey was designed. Four sets of questionnaires (5-point Likert scale) were developed for: patients/caregivers, physicians, managers, and technology providers. In this PoC phase, the survey demonstrates the feasibility of capturing divergent perspectives within a common monitoring logic, transforming systemic tensions into actionable data. Given the preliminary nature of the study, internal pre-testing was not conducted on a real-world sample but through Digital Twin Agents: synthetic profiles generated via AI. These agents were constructed to ensure a systematic coverage of the ecosystem's key configurations, specifically incorporating diverse age groups, varying levels of digital literacy and socio-economic backgrounds, distinct clinical conditions, different territorial contexts, and varying degrees of prior experience with telemedicine. These profiles were then used to complete the questionnaires, providing the data necessary to test the framework's analytical capacity. The resulting responses were transformed into synthetic indicators through a multi-level aggregation process. Items were mapped to a homogeneous scale and once normalized, aggregated following a hierarchical structure reflecting the framework's architecture. Initially, questionnaire items are grouped to generate scores for specific sub-dimensions; subsequently, these converge into the five main macro-dimensions. This bottom-up approach ensures full data traceability, allowing decision-makers to trace an overall score back to specific critical issues identified by a particular stakeholder category. The final step of this scoring procedure consists of synthesizing the results into a composite performance index. This TEMEF Overall index should not be interpreted as a simple arithmetic mean, but as a synthetic representation of the "health status" of the telemedicine system, so each domain score was computed by aggregating the questionnaire constructs associated with that specific dimension and rescaling the resulting value on a 1–100 scale. The TEMEF Overall Index was then calculated as the average of the five domain scores, thus providing a synthetic representation of the multidimensional performance profile generated by the simulated stakeholder responses. The strength of this logic lies in its ability to highlight systemic trade-offs: a high global index score may, in fact, conceal significant misalignments between underlying dimensions. In this sense, TEMEF scoring acts as a diagnostic signal, directing policy makers' attention toward domains where the gap between stakeholder expectations and operational results is most pronounced, thereby facilitating targeted, evidence-based governance interventions. Finally, the indicators were represented through a Power BI dashboard, conceived as an analytical environment to explore results by dimension, sub-dimension, and stakeholder group. As shown in Figure 2, this interface currently serves as a functional mockup and a demo decision support tool. At this stage, the dashboard does not constitute a continuous monitoring system but rather a preliminary interface to test the framework's usability and the clarity of the generated outputs.

Figure 2: Multidimensional Framework Validation Dashboard

4 Results and Discussion

The preliminary application of the TEMEF framework highlighted the potential of the proposed architecture to generate a multi-level and coherent representation of telemedicine system performance, supporting its operational feasibility as a synthesis tool for complex contexts, such as regional telemedicine systems. The results derive from simulated responses generated by synthetic stakeholder profiles and are therefore closely dependent on the assumptions embedded in the construction of those profiles. In this sense, TEMEF does not merely produce an average score; it offers a critical interpretative structure useful for exploring the distribution of value across different levels of analysis and among the various stakeholder involved. Even if, once the framework is applied to real-world populations, this scoring could be computed through a fuzzy logic methodology to better incorporate the inherent subjectivity and uncertainty of human perceptions, moving beyond the representative averages used for synthetic agents. This approach is aligned with the multidimensional logic of HTA, yet it integrates a multi-stakeholder perspective, allowing for the observation of interdependencies that traditional models often struggle to capture. The application produced an Overall Index of 65/100, suggesting a generally positive level of system maturity, although with relevant differences across domains. The clinical dimension emerged as the strongest area, with a score of 78, indicating that the platform appears particularly aligned with the objective of supporting care processes, clinical decision-making, and patient management. The economic dimension also showed a positive performance, with a score of 70, suggesting a favorable evaluation in terms of sustainability and value generation. The technological and organizational dimensions reported intermediate scores, respectively 66 and 65, pointing to an acceptable level of infrastructural and organizational readiness, while also indicating potential areas for refinement. By contrast, the experiential dimension obtained the lowest score, equal to 58, suggesting that user experience, usability, and digital literacy may represent critical issues for the full adoption of the platform. The results obtained should not be interpreted as static empirical evidence, but rather as a functional demonstration of the framework's ability to process heterogeneous data and deliver an analytical reading that ranges from the global composite index to the granularity of individual clinical, organizational, technological, experiential, and economic dimensions. The primary value emerged from this testing phase lies in the system's ability to make explicit the convergences and, more importantly, the divergences between the perspectives of the actors within the ecosystem. In the context of regional telemedicine, such divergences do not constitute mere informational noise; instead, they represent fundamental diagnostic signals of potential misalignments between strategic expectations and operational reality. By making these areas of friction visible, TEMEF enables policy makers to identify systemic trade-offs that a purely synthetic indicator would tend to obscure, such as the potential conflict between technological scalability and experiential acceptability for the end user. In this perspective, complexity is not isolated as a standalone variable but acts as a transversal logic permeating the entire framework. The five selected dimensions function as lenses through which to intercept crucial operational manifestations, such as stakeholder interdependence, the adaptive capacity of organizations, and the tensions between competing objectives. While the decision to limit the analysis to five macro-areas inevitably involves a degree of abstraction, this simplification proved functional for creating a tool that is both rigorous and flexible, capable of adapting to the institutional and territorial specificities of different regional contexts. The dashboard developed in Power BI completes this process, providing an analytical environment where the translation of theoretical constructs into clear visual outputs facilitates an informed, evidence-based decision-making process. However, it is essential to

reiterate the nature of this study as a PoC: the configurations analyzed derive from a testing phase conducted with synthetic profiles and should not be read as definitive results on the functioning of specific real-world services. The contribution lies in the proof of the model's operational translatability and its aptitude for supporting healthcare governance. Future developments must necessarily include empirical validation in the field to test the robustness of the indicators in real-world settings, potentially integrating TEMEF with more advanced dynamic modeling techniques, such as agent-based simulations, to better capture emerging phenomena. The ultimate goal remains the refinement of a tool capable of simulating stakeholder interactions and observing the evolution of trade-offs over time, transforming evaluation from a mere inspection act into strategic support for the sustainable innovation of local healthcare.

5 Ethics Declaration

This research did not require ethical approval, as it is based exclusively on publicly available secondary data and does not involve human participants, personal data, or sensitive information.

6 AI declaration

Artificial intelligence tools were used to support grammar and language refinement, item refinement, internal consistency checks, and the generation of synthetic stakeholder profiles and simulated responses for the preliminary pre-test. All outputs were critically reviewed and revised by the authors, who retain full responsibility for the study. AI was not used to collect empirical data or to replace empirical or psychometric validation.

References

Ashby, W. R. (1991). Requisite Variety and Its Implications for the Control of Complex Systems. In *Facets of Systems Science* (pp. 405–417). Springer US. https://doi.org/10.1007/978-1-4899-0718-9_28

Baby, B. E., & Chandran, G. J. (2022). *Digital Health Infrastructure* (pp. 701–714). https://doi.org/10.1007/978-981-16-6407-6_60

Baghbanian, A., Carter, D., & Merlin, T. (2025). Methods for the health technology assessment of complex interventions: A scoping review. *PLOS ONE*, *20*(3), e0315381. https://doi.org/10.1371/journal.pone.0315381

Begun JW, Zimmerman B, & Dooley KJ. (2003). Health care organizations as complex adaptive systems. In Mick SS & Wyttenbach ME (Eds.), *Advances in health care organization theory* (pp. 253–288). Jossey-Bass.

Benis, A., Tamburis, O., Chronaki, C., & Moen, A. (2021). One Digital Health: A Unified Framework for Future Health Ecosystems. *Journal of Medical Internet Research*, *23*(2), e22189. https://doi.org/10.2196/22189

Chang, H. (2015). Evaluation Framework for Telemedicine Using the Logical Framework Approach and a Fishbone Diagram. *Healthcare Informatics Research*, *21*(4), 230. https://doi.org/10.4258/hir.2015.21.4.230

Chiuchisan, I., Costin, H.-N., & Geman, O. (2014). Adopting the Internet of Things technologies in health care systems. *2014 International Conference and Exposition on Electrical and Power Engineering (EPE)*, 532–535. https://doi.org/10.1109/ICEPE.2014.6969965

Essen, A., Frishammar, J., & Cenamor, J. (2023). Entering non-platformized sectors: The Co-evolution of legitimacy debates and platform business models in digital health care. *Technovation*, *121*, 102597. https://doi.org/10.1016/j.technovation.2022.102597

Gleiss, A., Kohlhagen, M., & Pousttchi, K. (2021). An apple a day – how the platform economy impacts value creation in the healthcare market. *Electronic Markets*, *31*(4), 849–876. https://doi.org/10.1007/s12525-021-00467-2

Gonçalves-Bradley, D. C., Buckley, B. S., Fønhus, M. S., Glenton, C., Henschke, N., Lewin, S., Maayan, N., Mehl, G. L., Tamrat, T., & Shepperd, S. (2018). Mobile-based technologies to support client to healthcare provider communication and management of care. *Cochrane Database of Systematic Reviews*. https://doi.org/10.1002/14651858.CD012928

Greenhalgh, T., Wherton, J., Papoutsi, C., Lynch, J., Hughes, G., A'Court, C., Hinder, S., Fahy, N., Procter, R., & Shaw, S. (2017). Beyond Adoption: A New Framework for Theorizing and Evaluating Nonadoption, Abandonment, and Challenges to the Scale-Up, Spread, and Sustainability of Health and Care Technologies. *Journal of Medical Internet Research*, *19*(11), e367. https://doi.org/10.2196/jmir.8775

Kelly, C. J., & Young, A. J. (2017). Promoting innovation in healthcare. *Future Healthcare Journal*, *4*(2), 121–125. https://doi.org/10.7861/futurehosp.4-2-121

Khan, S., Vandermorris, A., Shepherd, J., Begun, J. W., Lanham, H. J., Uhl-Bien, M., & Berta, W. (2018). Embracing uncertainty, managing complexity: applying complexity thinking principles to transformation efforts in healthcare systems. *BMC Health Services Research*, *18*(1), 192. https://doi.org/10.1186/s12913-018-2994-0

Khodadad-Saryazdi, A. (2021). Exploring the telemedicine implementation challenges through the process innovation approach: A case study research in the French healthcare sector. *Technovation*, *107*, 102273. https://doi.org/10.1016/j.technovation.2021.102273

Kidholm, K., Ekeland, A. G., Jensen, L. K., Rasmussen, J., Pedersen, C. D., Bowes, A., Flottorp, S. A., & Bech, M. (2012). A MODEL FOR ASSESSMENT OF TELEMEDICINE APPLICATIONS: MAST. *International Journal of Technology Assessment in Health Care*, *28*(1), 44–51. https://doi.org/10.1017/S0266462311000638

Kim, K.-H., Kabir, E., & Ara Jahan, S. (2014). A Review of the Consequences of Global Climate Change on Human Health. *Journal of Environmental Science and Health, Part C*, *32*(3), 299–318. https://doi.org/10.1080/10590501.2014.941279

Konopik, J. (2023). The Impact of Digital Platforms and Ecosystems in Healthcare on Value Creation—A Integrative Review and Research Agenda. *IEEE Access*, *11*, 135811–135819. https://doi.org/10.1109/ACCESS.2023.3336983

Kristensen, F. B., Lampe, K., Wild, C., Cerbo, M., Goettsch, W., & Becla, L. (2017). The HTA Core Model ® —10 Years of Developing an International Framework to Share Multidimensional Value Assessment. *Value in Health*, *20*(2), 244–250. https://doi.org/10.1016/j.jval.2016.12.010

Li, X., Huang, L., Zhang, H., & Liang, Z. (2025). Enabling Telemedicine From the System-Level Perspective: Scoping Review. *Journal of Medical Internet Research*, *27*, e65932. https://doi.org/10.2196/65932

Plsek, P. E., & Greenhalgh, T. (2001). Complexity science: The challenge of complexity in health care. *BMJ*, *323*(7313), 625–628. https://doi.org/10.1136/bmj.323.7313.625

Rabbani, M. G., Alam, A., & Prybutok, V. R. (2025). Digital Health Transformation Through Telemedicine (2020–2025): Barriers, Facilitators, and Clinical Outcomes—A Systematic Review and Meta-Analysis. In *Encyclopedia* (Vol. 5, Number 4). Multidisciplinary Digital Publishing Institute (MDPI). https://doi.org/10.3390/encyclopedia5040206

Reis, J., Amorim, M., Melão, N., & Matos, P. (2018). *Digital Transformation: A Literature Review and Guidelines for Future Research* (pp. 411–421). https://doi.org/10.1007/978-3-319-77703-0_41

Ricci, G., Caraffa, A. M., & Gibelli, F. (2023). Telemedicine as a Strategic Tool to Enhance the Effectiveness of Care Processes: Technological and Regulatory Evolution over the Past Two Decades. In *Healthcare (Switzerland)* (Vol. 11, Number 5). MDPI. https://doi.org/10.3390/healthcare11050734

Rogers, D. G., Santamaria, K., Seng, E. K., & Grinberg, A. S. (2022). Behavioral Health, Telemedicine, and Opportunities for Improving Access. *Current Pain and Headache Reports*, *26*(12), 919–926. https://doi.org/10.1007/s11916-022-01096-w

Ross, J., Stevenson, F., Lau, R., & Murray, E. (2016). Factors that influence the implementation of e-health: A systematic review of systematic reviews (an update). In *Implementation Science* (Vol. 11, Number 1). BioMed Central Ltd. https://doi.org/10.1186/s13012-016-0510-7

Sood, S., Mbarika, V., Jugoo, S., Dookhy, R., Doarn, C. R., Prakash, N., & Merrell, R. C. (2007). What is telemedicine? A collection of 104 peer-reviewed perspectives and theoretical underpinnings. *Telemedicine and E-Health*, *13*(5), 573–590. https://doi.org/10.1089/tmj.2006.0073

Souza, V. de L. e, Oliveira, I. A. G. de, Caetano, R., Ribeiro, G. da R., Santos, D. L., Silva, A. B., & Silva, R. M. da. (2021). Conceptual frameworks used in the evaluation of Telehealth initiatives: A scoping review protocol. *Research, Society and Development*, *10*(6), e38910615913. https://doi.org/10.33448/rsd-v10i6.15913

Tabish, S. A., & Nabil, S. (2015). Future of healthcare delivery: Strategies that will reshape the healthcare industry landscape. *International Journal of Science and Research*, *4*(2), 727–758.

Talwar, S., Dhir, A., Islam, N., Kaur, P., & Almusharraf, A. (2023). Resistance of multiple stakeholders to e-health innovations: Integration of fundamental insights and guiding research paths. *Journal of Business Research*, *166*. https://doi.org/10.1016/j.jbusres.2023.114135

Yeganeh, H. (2019). An analysis of emerging trends and transformations in global healthcare. *International Journal of Health Governance*, *24*(2), 169–180. https://doi.org/10.1108/IJHG-02-2019-0012

Zhang, Y., Lin, Y.-Y., Lal, L. S., Reneker, J. C., Hinton, E. G., Chandra, S., & Swint, J. M. (2025). Stakeholder-driven multi-stage adaptive real-world theme-oriented (SMART) telehealth evaluation framework: a scoping review. *The Lancet Regional Health - Americas*, *44*, 101041. https://doi.org/10.1016/j.lana.2025.101041

Managing Knowledge for Including Persons with Disabilities in Organizational Contexts

Manuela Paolini[1], Domenico Raucci[2], Giuseppe Modarelli[3]
[1] University of L'Aquila, L'Aquila, Italy
[2] University "G. d'Annunzio" of Chieti-Pescara, Pescara, Italy
[3] University of Turin, Turin, Italy

Abstract

The 2030 Agenda of the United Nations promotes the inclusion of Persons with Disabilities (PwD) in the workplace as a core element of an organization's social responsibility (SR). Respecting this right to inclusion in the workplace is especially imperative for public sector organizations (PSOs), because of their institutional role and the intrinsic link between the public values they embody, such as solidarity, justice, equity, and social inclusion, and SR principles. However, merely implementing work accommodations, as defined by various international and national legislations, such as that of Italy, may not be enough to make PwD feel genuinely included in the workplace. To address these issues, knowledge of disability and inclusion principles must be carefully managed as an intangible resource to help employees recognize the importance of including PwD as an organizational SR value. Socialization mechanisms play a pivotal function in facilitating the transfer and internalization of knowledge within organizations. In these terms, informal controls (cultural and personnel controls) may represent a powerful tool for modelling and transferring knowledge, thereby affirming the principles of workplace inclusion for PwD. Based on self-determination theory, this study examined how the managerial use of informal controls experienced by non-executive employees in PSOs influences their prosocial behaviours towards PwD via the indirect effect of organizational identification. This prosocial outcome is particularly important in PSOs, as it represents an expression of the core role of public employees as public servants. To test the research hypotheses, data were collected by administering questionnaires to a sample of Italian municipal employees. Regression analysis verified all the hypothesized paths. This study has contributed to the emerging body of research on integrating SR principles into management control systems within PSOs. In doing so, it has also shed a light on the neglected role of informal controls as factors shaping and transferring the knowledge of inclusion values toward PwD to public employees.

Keywords: Knowledge management, Informal controls, Disability and inclusion principles, Public employees, Italy.

Paper type: Academic Research Paper

1 Introduction

The UN 2030 Agenda has promoted workplace inclusion (WI) of Persons with Disabilities (PwD) as a core element of organization's social responsibility (SR). Specifically, the European Employment Equality Directive and the European Disability Strategy 2021-2030 established the right of PwD to work on a fair and equal basis with other employees and to benefit from measures that support their integration into the workforce while avoiding any form of discrimination. Respecting this right to inclusion in the workplace is especially imperative for public sector organizations (PSOs), because of their institutional role and the intrinsic link between the public values they embody, such as solidarity, justice, equity, and social inclusion, and SR principles (Aguilera et al., 2007; Abdelmotaleb & Saha, 2019; Hammon, Gillis, and Icenogle, 2023). However, merely implementing work accommodations, variously defined by legislation in several countries, including Italy, may not be enough to make PwD feel genuinely included in the workplace (Shore et al., 2018; Waisman-Nitzan, Gal, and Schreuer, 2019; Coll & Mignonac, 2023), beyond limiting the effectiveness of disability management practices within PSOs (Palumbo, Hinna, and Manesh, 2023).

Knowledge, beliefs, and organizational culture around disability and inclusion need to be more deeply shaped so that PwD feel recognized, respected, and valued by colleagues for their diverse abilities (Randel et al., 2018; Mor-Barak et al., 2021).

To this end, PSOs should adopt approaches that dynamically nurture and manage knowledge, as an intangible resource, helping employees be aware of the relevance of PwD inclusion as an organizational SR value (Lee & Choi, 2003; Nelissen et al., 2017).

Socialization mechanisms play a pivotal function in facilitating knowledge transfer and internalization processes within organizations (Nonaka & Takeuchi, 1995; Pinho et al., 2012). Being embedded in organizations that encourage the progressive homogenization of values and goals increases their understanding, learning and

interiorization (Ravishankar and Pan, 2008; Kassa & King, 2023). These intangible processes of knowledge transfer make employees more psychologically empowered, thereby lending to behaviours aligned with organizational principles, goals, and expectations.

In PSOs contexts, promoting social-relational mechanisms aimed at spreading awareness of the value of inclusion among PwD, influencing the micro-foundations of employees' SR, could induce inclusive prosocial motivations and behaviours as an expression of their role as public servants (Grant and Gino, 2010; Abdelmotaleb & Saha, 2019).

The managerial use of informal controls, namely cultural and personnel controls, may serve these purposes by, respectively, building shared norms and values and inducing desired employee behaviour (Merchant and Van der Stede, 2007).

Research on the integration of SR principles within management control systems (Norris & O'Dwyer, 2004; Johnstone, 2019) suggests extending their traditional role of aligning employees' motivation and behaviour with the organization's expectations to support broader organizational implementation goals (Gond et al., 2012; Laguir, Laguir, and Tchemeni, 2019; Van der Kolk, Van Veen-Dirks, and ter Bogt, 2019). From this SR-oriented perspective, informal controls can be a powerful tool for modelling and transferring knowledge, fostering the affirmation of the principles of workplace inclusion for PwD.

At the individual level, the influences on employees resulting from managerial use of these controls can be usefully interpreted through the assumptions of self-determination theory (Ryan & Deci, 2000b; Gagné & Deci, 2005). According to this theory, when an individual engages with the organization and their need for belonging is satisfied, a series of positive effects are produced on the process of organizational identification (OI) and the internalization of organizational goals and values at the individual level. Effective managerial use of cultural and personnel controls may enable the alignment of employees' values and goals with those of the organization, by helping to fulfil the need for relatedness, that is, one's sense of belonging and significance within the reference community (Ryan & Deci, 2000b). Indeed, satisfying this need for consistency between personnel and organizational identity increases OI, promoting the internalization of PSOs' goals and values, including those related to the principles of disability and inclusion. Therefore, it is likely that the use of personnel and cultural controls, acting on the affective-relational dimensions of SR (Aguilera et al., 2007; D'Aprile & Talò, 2014), contributes to increasing employees' propensity to engage in prosocial behaviours towards PwD.

Previous studies have highlighted the links between prosocial behaviours and organizational inclusion (Ertas, 2014; Hammon, Gillis, and Icenogle, 2023; Nelissen et al., 2017), but few have investigated the role of management control systems in promoting inclusive principles and behaviours toward PwD (Tucker & Nash, 2023; Raucci & Paolini, 2025).

Building on this premise, the study explores the influence of managerial use of cultural and personnel controls, experienced by non-executive employees in PSOs, on their prosocial behaviours toward PwD via the indirect effect of OI.

This study contributed to the emerging research on integrating SR principles into management control systems (Norris & O'Dwyer, 2004; Gond et al., 2012; Johnstone, 2019) within PSOs. In particular, it shed light on the neglected role of cultural and personnel controls as factors shaping and transferring the knowledge of inclusion values toward PwD to public employees. This is a relevant topic in the context of PSOs because, by leveraging the micro-foundations of SR, it explores how perceptions of managerial use of informal controls can induce employees' inclusive prosocial behaviours toward PwD, as an expression of their nature as public servants (Bennett, 2011; Abdelmotaleb & Saha, 2019).

To test the research hypotheses, we collected data by administering questionnaires to a sample of employees in Italian municipalities.

The paper is structured as follows: Section 2 illustrates the theoretical framework and research hypotheses; Section 3 describes the sample and methodology; Section 4 presents the results; and Section 5 provides discussions and conclusions.

2 Background

PSOs should strengthen the integration of SR principles into their core values (Aguilera et al., 2007; Aguinis & Glavas, 2012), adopting management approaches that make employees aware of the relevance of PwD inclusion as an organizational value. Engaging employees in building and sharing the value of inclusion within the organizational context (Bennett, 2011) would allow them to develop a personal sense of SR (D'Aprile & Talò, 2014) and engage in inclusive prosocial behaviours toward colleagues with disabilities.

Building an inclusive work environment allows all internal stakeholders, especially those with disabilities, to experience a sense of belonging to their organizations, while feeling free to express their differences and

uniqueness (Metallo, Ricci, and Migliaccio, 2009; Migliaccio, 2019; Mor Barak, Luria, and Brimhall, 2021). Empirical evidence on the perceptions of inclusive work environments (Oware & Mallikarjunappa, 2021; Palumbo, Hinna, and Manesh, 2023) highlights improvements in the workplace well-being, attitudes, and behaviours of all employees, with broader positive effects on individual and organizational performance.

To achieve these goals, PSOs should foster and manage knowledge processes regarding the value of inclusion for the organization, to convey to employees not only formal notions of this principle, but also a deeper individual understanding of the ethical and SR implications of inclusion for PwD (Aguilera et al., 2007; Burns & Jollands, 2020; Pinho et al. 2012).

Indeed, knowledge management practices enable the collection, codification, sharing, and legitimization of knowledge related to disability, the rights of PwD, and the organizational value of inclusion, fostering a socio-organizational infrastructure capable of transforming institutional values into shared operational knowledge. In this regard, public management studies highlight the role of knowledge transfer and, in particular, knowledge socialization mechanisms, beyond those based on documents, in making knowledge sharing among employees effective (Nonaka & Takeuchi, 1995; Renaudin et al., 2018; Kassa & King, 2023).

When knowledge of the value of inclusion for PwD is effectively transferred to employees, fostering processes of internalizing this value, they learn that disability is not merely an individual issue, but the result of the interaction between the individual and organizational barriers (Butler & Ferlie, 2020; Yildiz et al., 2024). This changes employees' interpretative framework toward inclusion, which no longer appears as a discretionary concession, but as a collective responsibility of the organization to which socially responsible individual behaviours must be aligned. The knowledge thus internalized tends to translate into prosocial behaviours of inclusion toward PwD, such as helping, cooperating, adapting interactions, supporting accommodation, countering stereotypes, encouraging participation in decision-making, etc. (Grant and Gino, 2010; Sanclemente et al., 2024). Indeed, these inclusive prosocial behaviours emerge more easily when the employee has both knowledge and signals from the organization that legitimize such actions (Nelissen et al., 2017; Raucci & Paolini, 2025).

These cognitive purposes regarding the value of inclusion for the organization can be incorporated into the traditional purposes of using management control systems to pursue the integration of SR objectives into the functions of these systems, as suggested by recent studies (Gond et al., 2012; Johnstone, 2019; Laguir, Laguir, and Tchemeni, 2019). In this sense, management control systems would support internal knowledge management processes regarding the value of inclusion for PwD and increase the congruence of employee goals and behaviours with those of the organization (Norris & O'Dwyer, 2004; Tessier & Otley, 2012; Van der Kolk, Van Veen-Dirks, and ter Bogt, 2019).

Adopting Merchant and Van der Stede's (2007) conceptual model, management control systems can function as formal mechanisms (such as results and action controls) and informal mechanisms (such as personnel or culture controls) to induce alignment between individual and organizational goals. Informal controls, operating according to the dynamics of social control, focus attention on employees' values and behavioural dimensions. Personnel controls foster behavioural congruence with organizational expectations by employing human management practices. Cultural controls aim to shape, communicate, and disseminate the organization's core values, principles, and norms among employees. Several studies have highlighted that these informal controls, if appropriately employed in management control systems, can be as effective as formal controls in influencing organizational performance (Laguir, Laguir, and Tchemeni, 2019; Goebel & Weißenberger, 2017).

The interface between management control systems and micro-SR outcomes is an emerging topic in the PSOs context. In these organizations, which have complex objectives and responsibilities related to public value (Abdelmotaleb and Saha, 2019; Bracci et al., 2019; Burns & Jollands, 2020), the attitudinal and behavioural potential of cultural and personnel controls can be crucial as a tool for transferring knowledge of the value of inclusion to PwD. Previous studies have highlighted the links between organizational inclusion and prosocial behaviours driven by individual perceptions of SR (Aguinis and Glavas, 2012; Ertas, 2014; Nelissen et al., 2017; Hammon, Gillis, and Icenogle, 2023). However, few studies have explored the role of management control systems, and, more specifically, informal controls, in promoting principles of organizational inclusion toward PwD (Tucker & Nash, 2023; Raucci & Paolini, 2025).

The use of personnel and cultural control as drivers of employee micro-SR can be explored by adopting self-determination theory (Ryan & Deci, 2000b). This theory posits that individuals aim to satisfy fundamental needs such as competence, autonomy, and relatedness. This latter need is satisfied when employees interact with internal organisational stakeholders. A sense of relatedness facilitates belonging and individuals' identification with organizational goals and values (Gagné & Deci, 2005). By satisfying employees' need for belonging, cultural and personnel controls enable them to share the organization's values and transfer knowledge into employees' interpretative frameworks. This process deepens their emotional attachment and emotional bond with the

organization, fostering identification with the SR goals and values conveyed by the PSO (Aguilera et al., 2007; D'Aprile & Talò, 2014). OI dynamics (Abdelmotaleb & Saha, 2019; Coll & Mignonac, 2023; Smidts et al., 2001) foster employees' perceptions that expectations and norms exist that are consistent with their personal identity and that they should actively participate in the pursuit of the organization's SR values. Therefore, it is likely that, through these micro-SR influences, personnel and cultural controls increase employees' propensity to align with the inclusive values conveyed by their PSOs, activating inclusive prosocial behaviours toward PwD.

Based on this rationale, we hypothesize that:

- *Hp 1: Cultural control positively influences organizational identification of public employees.*
- *Hp 2: Personnel control positively influences organizational identification of public employees.*
- *Hp 3: Organizational identification positively influences prosocial behaviours toward PwD among public employees.*

3 Methods

An empirical analysis was conducted on a sample of Italian municipalities. We focused on municipalities from the ISTAT demographic range of 20,000- 59,999 residents, which is representative of a significant proportion of Italian municipalities. Then, we selected those municipal entities with over 50,000 residents and at least 200 staff members, as, based on the Italian regulation (2009 'White Paper on Accessibility and Urban Mobility', L.D. 75/2017 and L. 227/2021), these municipalities are required to implement disability management practices and appoint a figure responsible for them. This selection process allowed to identify 19 municipalities, of which 9 agreed to participate in the survey.

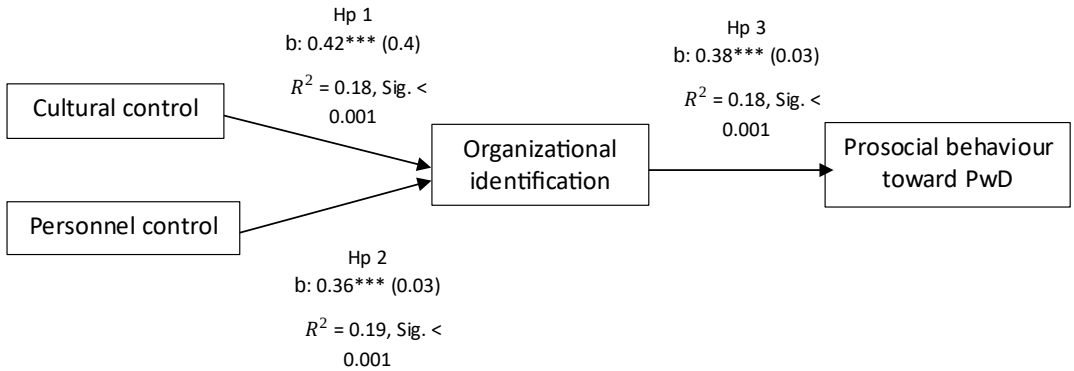
A survey was conducted to collect data from non-executive employees, assuring them privacy and confidentiality. Cultural and personnel controls were each measured using 5 items based on Merchant and Van der Stede (2007), as adopted by Van der Kolk, Van Veen-Dirks, and ter Bogt (2019) (a sample items includes "Employees know what is expected from them from the prevailing culture" for cultural control and "'Many opportunities for training and education are being offered" for personnel control). OI was measured using 5 items from Smidts et al.'s (2001) scale, which assess employees' emotional attachment and affective bond to their organization through items such as "I experience a strong sense of belonging to my organization". Prosocial behaviour towards PwD was evaluated using a 4-items scale developed by Grant (2008), who assess this behaviour by measuring the reasons beyond the motivation to work (e.g., "I am motivated to do my work because I care about benefiting others through my work").

The final sample included 411 employees. Preliminary analyses were conducted to evaluate the psychometric soundness of the measures. Internal consistency was assessed using Cronbach's alpha coefficients, and exploratory factor analyses were employed to verify the scales' convergent and discriminant validity. Descriptive statistics and bivariate correlations were computed for all variables. The model was tested using a linear regression.

4 Findings

The hypotheses testing, which resulted from regression analysis, showed that cultural and personnel controls had a positive significant effect on OI ($\beta = 0.42$, $p \leq 0.001$ and $\beta = 0.36$, $p \leq 0.001$, respectively). This result is consistent with the assumption that the managerial uses of informal controls contribute to satisfying employees' need for belonging and build organizational identity.

Figure 1 represents the evidence.



Standard error in brackets / ***p-value 0.001

Figure 1: Regression analysis

The findings also revealed a positive and significant relationship between OI and prosocial behaviour towards PwD (b = 0.38, p £ 0.001). Thus, strengthening employees' ties with the organization and its goals allows for a substantial increase in the WI of PwD. To this end, results highlighted the importance of reinforcing the relational dimension induced by informal controls as a possible means of knowledge management and transfer, capable of reinforcing the relations of public servants with the public values and goals of the organization where they work. Thus, all the paths hypothesized were verified.

5 Discussion and Conclusions

This study investigated whether and how informal controls—specifically cultural and personnel controls—support prosocial behaviours toward PwD within PSOs. To address this aim, we conducted an empirical survey-based analysis involving non-executive employees working in Italian municipalities.

Overall, the results aligned with the mechanisms posited by self-determination theory (Gagné & Deci, 2005). In particular, both cultural and personnel controls appeared to foster congruence between public employees' values and objectives and those promoted by their organization, thereby strengthening OI. This evidence is consistent with earlier work showing that cultural and personnel controls can satisfy employees' need for relatedness (Van der Kolk, Van Veen-Dirks, and ter Bogt), facilitating the circulation and internalization of shared organizational values. In PSOs, such values often include SR principles, whose foundations are interwoven with public value and its multiple institutional roots (Abdelmotaleb and Saha, 2019; Bracci et al., 2019). In addition, the findings indicated that stronger OI, reflecting a more intense affective attachment to the organization and to the goals pursued by public servants, reinforced employees' relationship with the public service mission by conveying more prosocial behaviour toward PwD. Taken together, the empirical evidence supported all the hypothesized paths.

By doing so, the study added to existing knowledge on the motivational implications of how informal elements of management control systems are used, focusing specifically on public employees who do not hold managerial responsibilities. Building on this, the study contributed to emerging debates on connecting management control systems with the micro-foundations of SR (Johnstone, 2019) in PSOs (Aguinis and Glavas, 2012; Hammon, Gillis, and Icenogle, 2023). More specifically, the analysis brings attention to the often-overlooked role of cultural and personnel controls in shaping and transmitting PwD inclusion-related knowledge and values to public employees. It also highlighted the importance of employees' perceived experience of control, which remains underexplored in this stream of research (Tessier and Otley, 2012). In this way, the paper informed the growing literature on SR micro-outcomes (Aguilera et al., 2007; Aguinis and Glavas, 2012; Hammon, Gillis, and Icenogle, 2023) within the still relatively under-investigated setting of PSOs (Raucci and Paolini, 2025).

From a practical standpoint, public management should embed social inclusion as an internally driven principle rather than treating it primarily as a response to external requirements. An internal orientation may enable organizations to mobilize diverse knowledge drivers to build a culture that extends beyond workplace adjustments and more fully reflects the SR principles that characterize PSOs. In this respect, public managers should aim to integrate and coordinate personnel and cultural controls more closely with human resource management practices, which are directly implicated in disability management.

Promoting inclusion "from within" through tighter alignment between personnel and cultural controls and human management practices may support the co-construction of sustainable development pathways, helping PSOs to pursue the wider public values and objectives they are expected to embody. Encouraging inclusive prosocial behaviours toward PwD may facilitate the effective implementation of disability management practices, moving beyond a narrow focus on accommodations toward a deeper cultural orientation grounded in SR. Such behaviours are likely to yield benefits for PwD (performance, well-being, and turnover intentions) and for service relationships with PwD as external stakeholders of public services.

Notwithstanding these insights, the research was subject to limitations. The evidence relies on survey data and a relatively small set of observations, which constrains generalizability because responses may reflect the specific organizational context and respondents' subjective interpretations. Future research should expand the sample to verify whether the observed relationships hold across other municipalities and across different types of PSOs. Moreover, subsequent studies should examine the role of additional intervening factors that may shape these dynamics.

6 Ethics declaration

Ethical approval was not necessary for this study.

7 AI declaration

The Authors declare that no AI tool was used in developing the paper.

References

Abdelmotaleb, M., and Saha, S.K. (2019) Corporate social responsibility, public service motivation and organizational citizenship behavior in the public sector, *International Journal of Public Administration*, Vol. 42, No. 11, pp. 929–939.

Aguilera, R.V., Rupp, D.E., Williams, C.A., and Ganapathi, J. (2007) Putting the S back in corporate social responsibility: A multilevel theory of social change in organizations, *Academy of Management Review*, Vol. 32, No. 3, pp. 836–863.

Aguinis, H., and Glavas, A. (2012) What we know and don't know about corporate social responsibility: A review and research agenda, *Journal of Management*, Vol. 38, pp. 932–968.

Bennett, A.J.W. (2011) Learning to be job ready: Strategies for greater social inclusion in public sector employment, *Journal of Business Ethics*, Vol. 104, No. 3, pp. 347-359.

Bracci, E., Papi, L., Bigoni, M., Deidda Gagliardo, E., and Bruns, H.J. (2019) Public value and public sector accounting research: A structured literature review, *Journal of Public Budgeting, Accounting & Financial Management*, Vol. 31, No. 1, pp. 103–136.

Burns, J., and Jollands, S. (2020) Acting in the public interest: Accounting for the vulnerable, *Accounting and Business Research*, Vol. 50, No. 5, pp. 507–534.

Butler, M.J., and Ferlie, E. (2020) Developing absorptive capacity theory for public service organizations: Emerging UK empirical evidence, *British Journal of Management*, Vol. 31, No. 2, pp. 344-364.

Coll, C., and Mignonac, K. (2023) Perceived organizational support and task performance of employees with disabilities: A need satisfaction and social identity perspectives, *The International Journal of Human Resource Management*, Vol. 34, No. 10, pp. 2039–2073.

D'Aprile, G., and Talò, C. (2014) Measuring corporate social responsibility as a psychosocial construct: A new multidimensional scale, *Employee Responsibilities & Rights Journal*, Vol. 26, No. 3, pp. 153–175.

Ertas, N. (2014) Public service motivation theory and voluntary organizations: Do government employees volunteer more? *Nonprofit and Public Sector Quarterly*, Vol. 43, No. 2, pp. 254–271.

Gagné, M., and Deci, E.L. (2005) Self-determination theory and work motivation, *Journal of Organizational Behavior*, Vol. 26, No. 4, pp. 331–362.

Goebel, S., and Weißenberger, B.E. (2017) The relationship between informal controls, ethical work climates, and organizational performance, *Journal of business ethics*, Vol. 141, No. 3, pp. 505-528.

Gond, J.P., Grubnic, S., Herzig, C., and Moon, J. (2012) Configuring management control systems: Theorizing the integration of strategy and sustainability, *Management accounting research*, Vol. 23, No. 3, pp. 205-223.

Grant, A.M. (2008) Does intrinsic motivation fuel the prosocial fire? Motivational synergy in predicting persistence, performance, and productivity, *Journal of applied psychology*, Vol. 93, No. 1, pp. 48.

Grant, A.M., and Gino, F. (2010) A little thanks goes a long way: Explaining why gratitude expressions motivate prosocial behavior, *Journal of personality and social psychology*, Vol. 98, No. 6, pp. 946.

Hammon, S.D., Gillis, W.E., and Icenogle, M.L. (2023) From social responsibility to employee engagement: Evidence from the public sector, *Journal of Management and Governance*, Vol. 27, No. 4, pp. 1223–1247.

Johnstone, L. (2019) Theorising and conceptualising the sustainability control system for effective sustainability management, *Journal of Management Control*, Vol. 30, No. 1, pp. 25–64.

Kassa, E.T., and Ning, J. (2023) A systematic review on the roles of knowledge management in public sectors: Synthesis and way forwards, *Heliyon*, Vol. 9, No. 11.

Laguir, L., Laguir, I., and Tchemeni, E. (2019) Implementing CSR activities through management control systems: A formal and informal control perspective, *Accounting, Auditing & Accountability Journal*, Vol. 32, No. 2, pp. 531–555.

Lee, H. and Choi, B. (2003) Knowledge management enablers, processes, and organizational performance: an integrative view and empirical examination, *Journal of Management Information Systems*, Vol. 20, No. 1, pp. 179-228.

Merchant, K.A., and Van der Stede, W.A. (2007) *Management control systems: Performance measurement, evaluation and incentives*, Prentice Hall.

Metallo, G., Ricci, P., and Migliaccio, G. (Eds.) (2009) *La risorsa umana "diversamente abile" nell'economia dell'azienda*, Giappichelli.

Migliaccio, G. (2019) Disabled people in the stakeholder theory: A literature analysis, *Journal of the Knowledge Economy*, Vol. 10, No. 4, pp. 1657–1678.

Mor Barak, M.E., Luria, G., and Brimhall, K.C. (2021) What leaders say versus what they do: Inclusive leadership, policy-practice decoupling, and the anomaly of climate for inclusion, *Group & Organization Management*, Vol. 47, pp. 840–871.

Nelissen, P.T.J.H., Hülsheger, U.R., Van Ruitenbeek, G.M.C., and Zijlstra, F.R.H. (2017) Lending a helping hand at work: A multilevel investigation of prosocial motivation, inclusive climate and inclusive behavior, *Journal of Occupational Rehabilitation*, Vol. 27, pp. 467–476.

Nonaka, I., and Takeuchi, H. (1995) *The Knowledge-Creating Company*, Oxford University Press.

Norris, G., and O'Dwyer, B. (2004) Motivating socially responsive decision making: the operation of management controls in a socially responsive organisation, *The British Accounting Review*, Vol. 36, No. 2, pp. 173-196.

Oware, K.M., and Mallikarjunappa, T. (2021) Disability employment and financial performance: The effect of technological innovation of listed firms in India, *Social Responsibility Journal*, Vol. 17, No. 3, pp. 384–398.

Palumbo, R., Hinna, A., and Manesh, M.F. (2023) Aiming at inclusive workplaces: A bibliometric and interpretive review at the crossroads of disability management and human resource management, *Journal of Management & Organization*, pp. 1-24.

Pinho, I., Rego, A., Pina and Cunha, M. (2012) Improving knowledge management processes: a hybrid positive approach, *Journal of knowledge management*, Vol. 16 No. 2, pp. 215-242.

Randel, A.E., Galvin, B.M., Shore, L.M., Ehrhart, K.H., Chung, B.G., Dean, M.A., and Kedharnath, U. (2018) Inclusive leadership: Realizing positive outcomes through belongingness and being valued for uniqueness, *Human resource management review*, Vol. 28, No. 2, pp. 190-203.

Raucci, D., and Paolini, M. (2025) *How Do Cultural and Personnel Controls Influence Inclusive Workplace Climate for People with Disability? Insights from the Italian Context*. In Diversity and Equity in Accounting: Emerging Issues, Challenges and Opportunities, pp. 189-201. Cham: Springer Nature Switzerland.

Ravishankar, M.N., and Pan, S.L. (2008) The influence of organizational identification on organizational knowledge management (KM), *Omega*, Vol. 36, No. 2, pp. 221-234.

Renaudin, M., Dal Mas, F., Garlatti, A., and Massaro, M. (2018) Knowledge Management and cultural change in a knowledge intensive public organization. The case of Swiss medic, *4th Knowledge Management and Intellectual Capital Excellence Awards*, pp. 85-96.

Ryan, R.M., and Deci, E.L. (2000b) Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being, *American Psychologist*, Vol. 55, No. 1, pp. 68–78.

Sanclemente, F.J., Gamero, N., Medina, F.J., and Mendoza-Denton, R. (2024) A multilevel model of job inclusion of employees with disabilities: The role of organizational socialization tactics, coworkers social support, and an inclusive team context, *Applied Psychology*, Vol. 73, No. 3, pp. 887-909.

Shore, L.M., Cleveland, J.N., and Sanchez, D. (2018) Inclusive workplaces: A review and model, *Human Resource Management Review*, Vol. 28, No. 2, pp. 176-189.

Smidts, A., Pruyn, A.T.H., and Van Riel, C.B. (2001) The impact of employee communication and perceived external prestige on organizational identification, *Academy of Management Journal*, Vol. 44, No. 5, pp. 1051-1062.

Tessier, S., and Otley, D. (2012) A conceptual development of Simons' levers of control framework, *Management Accounting Research*, Vol. 23, No. 3, pp. 171–185.

Tucker, B.P., and Nash, E. (2023) Good for business: Towards a management control research agenda relating to the employment of workers with intellectual disability, *Journal of Accounting Literature*.

Van der Kolk, B., Van Veen-Dirks, P. M., and ter Bogt, H. J. (2019) The impact of management control on employee motivation and performance in the public sector, *European Accounting Review*, Vol. 28, No. 5, pp. 901–928.

Waisman-Nitzan, M., Gal, E., and Schreuer, N. (2019) Employers' perspectives regarding reasonable accommodations for employees with autism spectrum disorder, *Journal of Management & Organization*, Vol. 25, No. 4, pp. 481-498.

Yildiz, H.E., Murtic, A., Morgulis-Yakushev, S., and Klofsten, M. (2024) Individual-level absorptive capacity: Unveiling the interplay between dispositions and work context, *Technovation*, Vol. 131, pp. 102965.

Absorptive Capacity in Complex Public Organizations: The Role of AI

Silvia Amélia Fonseca de Oliveira, Nathália Berger Werlang
Universidade Federal de Santa Catarina, Florianópolis, Brasil

Abstract

Artificial Intelligence has emerged as a relevant enabler of Knowledge Management, particularly in complex public organizations exposed to large volumes of external information and increasing demands for innovation and public value creation. In such contexts, international missions constitute strategic channels for accessing external knowledge through cooperation networks, foreign institutions, and professional interactions. However, while organizations may effectively identify and acquire valuable knowledge through these initiatives, significant challenges often remain in converting it into institutional learning and sustained organizational outcomes. Accordingly, the objective of this article is to understand how AI can act as a facilitating element of Absorptive Capacity in public organizations characterized by high organizational complexity and intense cooperation with international networks. The study adopts a qualitative single-case study design conducted in a federal public organization. Empirical evidence was generated through documentary analysis and 18 semi-structured interviews with public servants assigned to international missions in different countries and international organizations. Data were analyzed through content analysis supported by Atlas.ti, using the refined AC model proposed by Todorova and Durisin (2007) as the analytical framework. The findings indicate that AC is manifested in the context of international missions, particularly in the stages of recognition of value, knowledge acquisition and assimilation. Public servants assigned abroad perform functions analogous to gatekeepers, connecting the organization to strategic knowledge sources and international networks. Prior knowledge, technical expertise, and relational embeddedness were identified as critical antecedents supporting these early absorptive processes. However, significant bottlenecks emerge in the later stages of transformation and exploitation. The results reveal weaknesses in knowledge diffusion routines, fragmented feedback cycles, dependence on individual actors, and institutional memory losses associated with personnel turnover. Based on these findings, the article argues that AI may strengthen organizational learning processes through intelligent scanning, information filtering, knowledge repositories, pattern recognition, and decision-support systems. Such mechanisms may help public organizations reduce information overload, improve knowledge circulation, and convert externally acquired knowledge into innovation and sustained public value. Nevertheless, AI should complement rather than replace human judgment, contextual interpretation, and relational capabilities.

Keywords: Absorptive Capacity; Artificial Intelligence; Knowledge Management; Public Organizations; International Missions.

Paper type: Practical Paper

1 Introduction

Contemporary public organizations increasingly operate in complex, networked, and interorganizational environments characterized by information overload, institutional fragmentation, and growing expectations regarding the creation of public value. In such contexts, external knowledge generated through international cooperation and cross-boundary interactions often holds significant strategic potential. However, public organizations frequently struggle to systematically internalize, disseminate, and apply the knowledge acquired through these initiatives. This challenge highlights the relevance of Absorptive Capacity (AC), understood as an organizational capability to recognize the value of external knowledge, acquire it, assimilate it, transform it, and apply it (Cohen and Levinthal, 1990; Todorova and Durisin, 2007; Zahra and George, 2002).

In the public sector, AC assumes particular importance due to governance complexity, bureaucratic structures, and accountability demands (Ajmal et al., 2025; Fagundes et al., 2025; Harvey et al., 2010; van der Graaf et al., 2021). Although public organizations are increasingly exposed to valuable external knowledge sources, prior studies indicate persistent difficulties in converting such knowledge into institutionalized practices, innovation, and improved service delivery (Elbashir et al., 2022; Pinho et al., 2024; Waheed, 2025). Empirical evidence also suggests that while acquisition and assimilation processes are more visible, transformation and exploitation often remain constrained by fragmentation, rigid routines, and limited coordination (Ajmal et al., 2025; Van der Graaf et al., 2021; Castro et al., 2022; Pinho et al., 2024; Waheed, 2025).

These challenges become particularly salient in the context of international missions, which can function as strategic arenas for organizational learning by providing access to external knowledge and opportunities for collaboration (Françoso and Vonortas, 2023; Jin and Wang, 2021; Mori et al., 2025). Nevertheless, the knowledge generated in such initiatives is often dispersed across individuals, weakly codified, or insufficiently shared internally (Murray et al., 2011; Stark, 2018; van der Graaf et al., 2021; Zhou et al., 2021).

At the same time, recent developments in Knowledge Management (KM) have emphasized the growing potential of Artificial Intelligence (AI) to support information filtering, semantic analysis, knowledge retrieval, and decision support (Su et al., 2022). AI may enhance organizational capabilities to process large volumes of heterogeneous information and transform dispersed inputs into actionable knowledge (Mohan, 2025; Zhou, 2021). From this perspective, AI emerges as a promising enabling mechanism for strengthening absorptive capacity processes.

Despite growing interest in both AC and AI-enabled knowledge management, limited attention has been given to how AI can support specific AC processes in public organizations. Based on empirical evidence from international missions: **how can Artificial Intelligence facilitate the absorption of external knowledge in complex public organizations**?

Accordingly, the objective of this article is to understand how AI can act as a facilitating element of Absorptive Capacity in public organizations characterized by high organizational complexity and intense cooperation with international networks.

Methodologically, this study adopts a qualitative research design grounded in the AC model proposed by Todorova and Durisin (2007). Empirical findings from international missions are used to identify recurrent barriers and enabling conditions associated with absorptive processes. Building on these findings, the article advances a conceptual discussion on how AI tools and practices may help overcome such barriers.

This article contributes to the literature in three ways. First, it extends AC research to the context of public-sector international missions as strategic learning environments. Second, it connects empirical evidence on AC barriers with emerging debates on AI-enabled knowledge management. Third, it offers practical insights for managers seeking to improve the conversion of externally acquired knowledge into public value through more intelligent and systematic organizational processes.

2 Theoretical background

This section discusses AC, its relevance in public organizations, and the potential role of AI in supporting knowledge processes.

2.1 Absorptive Capacity as a Dynamic Capability

The concept of AC was introduced by Cohen and Levinthal (1989, 1990) to describe an organization's ability to recognize the value of external knowledge, assimilate it, and apply it for productive purposes. Their perspective establishes AC as a cumulative learning capability rooted in prior related knowledge, emphasizing that existing knowledge shapes the organization's ability to identify and integrate new information.

Building on this foundation, Zahra and George (2002) distinguish between Potential Absorptive Capacity (PACAP), composed of acquisition and assimilation, and Realized Absorptive Capacity (RACAP), related to transformation and exploitation. Their contribution highlights that access to knowledge does not automatically lead to productive use, making social integration mechanisms essential for reducing the gap between potential and realized value.

Todorova and Durisin (2007) revisit this model by reintroducing the stage of recognition of value and conceptualizing AC as a recursive process characterized by feedback loops among stages. They also emphasize the relevance of power relationships, which may act as filters or barriers affecting the assimilation, transformation, and use of external knowledge.

The operationalization of AC often depends on gatekeepers, individuals positioned between internal systems and external knowledge sources (Cohen and Levinthal, 1990). These actors play a central role in environmental scanning, knowledge acquisition, and translating information into organizationally meaningful forms (Murray et al., 2011). In international contexts, they also mediate norms, practices, and sensitive information across institutional boundaries (Françoso and Vonortas, 2022). By maintaining ties with both external partners and internal structures (Graf, 2010), gatekeepers help reduce communication gaps and cognitive incompatibilities (Daghfous, 2004).

These theoretical developments suggest that AC is particularly relevant in environments characterized by complexity, interdependence, and continuous interaction with heterogeneous external knowledge sources.

2.2 Absorptive Capacity in Public Organizations

Although originally developed in private-sector settings, AC has become increasingly relevant in Public Administration, where organizations face institutional complexity, multiple stakeholders, bureaucratic

constraints, and growing expectations regarding efficiency, innovation, and public value creation. In such contexts, the ability to identify, internalize, and apply external knowledge becomes strategically important.

Elbashir et al. (2022) demonstrate that the assimilation of business intelligence systems in public organizations depends substantially on the effective incorporation of external knowledge. In this sense, AC may function as a compensatory mechanism for structural deficiencies, such as limitations in talent retention, thereby enabling more robust data-driven decision-making processes.

Recent studies have also challenged the linearity traditionally associated with AC models. Butler and Ferlie (2020) argue that, in public organizations, knowledge acquisition unfolds as a continuous, relational, and co-creative dynamic rather than through discrete sequential phases. Similarly, Fagundes, Tumelero and Fagundes (2025) introduce the concept of AC metaroutines, suggesting that knowledge absorption occurs through recursive and iterative sequences of organizational practices.

Taken together, these findings indicate that AC in Public Administration should be understood not merely as a set of organizational capabilities, but as an emergent, path-dependent, and socially embedded process whose effectiveness depends on conditions that enable the transformation of knowledge into sustained public value.

2.3 Artificial Intelligence in Knowledge Management

Recent debates in Knowledge Management have emphasized the growing relevance of AI as a potential enabler of processes that transform large volumes of heterogeneous information into organizationally useful knowledge (Zhou, 2021; Mohan, 2025). In knowledge-intensive environments, AI may support information filtering, pattern recognition, semantic analysis, summarization, and decision support.

Prior studies also suggest that digital capabilities may reinforce AC. Joshi et al. (2010) mobilized AC theory to explain how information technology tools strengthen knowledge capabilities, while Mao et al. (2017) and Ma et al. (2021) indicate that technological competence and IT infrastructure may operate in synergy with AC to promote agility, performance, and adaptive capacity.

From this perspective, AI may contribute to specific dimensions of AC. In recognition and acquisition stages, AI can support environmental scanning and automated retrieval of relevant information. During assimilation, AI tools may facilitate classification, summarization, and interpretation of dispersed content. In transformation processes, AI may help identify patterns, combine heterogeneous knowledge sources, and reveal actionable insights. In exploitation, AI-based systems may support decision-making and implementation processes.

These contributions are particularly relevant in public organizations, which often face information fragmentation, interdepartmental silos, and what Zhou, Chen and Han (2021) describe as data stickiness. Under such conditions, AI may help reduce analytical overload, improve knowledge flows, and strengthen organizational learning. However, the adoption of AI in public settings also requires governance, transparency, data quality, and human oversight. Its value depends not only on technological availability, but also on the organization's preexisting AC to interpret, adapt, and apply digitally supported knowledge processes.

3 Methodology

This study adopts a qualitative single-case study design with an exploratory and descriptive orientation. The empirical setting is a federal public organization whose international missions constitute strategic arenas for external knowledge acquisition, interorganizational interaction, and organizational learning. A qualitative approach was considered appropriate because it enables an in-depth understanding of complex organizational phenomena, particularly those involving tacit knowledge flows and learning processes across organizational boundaries.

Three complementary data collection techniques were employed. First, a bibliographic review was conducted using databases such as Scielo, Web of Science, and Scopus to consolidate the theoretical background and support the selection of the AC framework. Second, documentary research involved the analysis of regulations, institutional documents, and digital sources related to the organization's international activities. Third, and most centrally, semi-structured interviews were conducted with public servants directly involved in missions abroad.

The empirical corpus comprises 18 interviews with public servants assigned to international missions in different countries and international organizations. A semi-structured interview protocol was developed to ensure comparability while preserving flexibility for emerging issues. The protocol was organized into four sections aligned with the research objectives: (1) prior knowledge and external knowledge sources; (2) routines and processes related to recognition of value, acquisition, assimilation, transformation, and application of external knowledge; (3) contingent factors such as social integration, power relationships, feedback mechanisms, and internal/external interactions; and (4) suggestions for organizational improvement.

Data were analyzed through content analysis following Bardin (2021), comprising pre-analysis, material exploration, and treatment of results through inference and interpretation. Coding procedures were initially guided by predefined categories derived from the refined AC model proposed by Todorova and Durisin (2007), while remaining open to emergent categories arising from the empirical material. Atlas.ti (version 25) was used to support data organization and coding procedures.

The methodological strategy unfolded in two complementary stages. First, the qualitative case study examined how AC processes manifest in international missions, identifying recurrent barriers, enabling conditions, and organizational routines related to external knowledge use. Second, these findings were conceptually reinterpreted in light of recent literature on AI in KM. Rather than treating AI as an empirically observed phenomenon, the study uses the literature to explore how AI-enabled mechanisms can strengthen AC in public organizations operating in international contexts.

4 Results and Discussion

The empirical analysis of international missions in a public organization indicates that AC is clearly manifested in the public sector. Public servants operating abroad perform key gatekeeper functions in mediating external knowledge flows. The findings suggest that early absorptive stages, particularly recognition of value, acquisition, and initial assimilation, tend to operate more effectively than later stages.

4.1 Absorptive Capacity Dynamics in International Missions

The empirical findings indicate that international missions constitute strategic organizational environments in which AC becomes clearly observable in practice. In these settings, public servants assigned abroad operate at the interface between the organization and external institutional environments, facilitating access to new information, cooperation opportunities, operational practices, and international knowledge sources. This evidence reinforces the relevance of AC in public-sector contexts characterized by interorganizational interaction and external dependence.

A first relevant finding concerns the antecedents of AC. The study identified a wide diversity of external knowledge sources accessed during missions abroad, varying according to the profile and specific responsibilities of each mission. These sources include formal institutional channels, professional networks, interpersonal relationships, and situational interactions developed in the host environment. Access to these sources is described by interviewee 10, when he mentions "*meetings with local authorities, participation in events and seminars, technical visits to local institutions, in addition to direct contact with attachés and liaison officers from other countries*". Such diversity confirms the importance of external knowledge exposure as a precondition for absorptive processes, while also indicating that different international positions may require differentiated KM strategies.

The results further suggest that prior knowledge held by public servants plays a decisive role in enabling absorptive dynamics. Accumulated experience within the organization was shown to be particularly relevant for recognizing the value of external knowledge, especially in contexts where preparation processes for missions were limited or operational guidance was insufficient. Prior technical expertise, familiarity with organizational routines, and previous experience in related areas enhanced the ability of individuals to interpret external signals and identify strategically relevant information. These findings are consistent with Cohen and Levinthal (1990) and Todorova and Durisin (2007), who emphasize prior knowledge as a critical antecedent of AC.

Regarding the core dimensions of AC, the recognition of value emerged as strongly dependent on the profile and competencies of the civil servant assigned to the mission. The ability to identify relevant information was influenced by technical background, contextual understanding, and relational positioning within international environments. In this sense, the recognition stage does not appear automatic, but rather contingent upon individual and contextual conditions, as mentioned by interviewee E4, when stating that "*I would not say that the management guided me much, because what guided me was my own experience*".

Similarly, knowledge acquisition was described as a recurrent and relatively effective process, supported by institutional embeddedness and broad social networks. Access to relevant knowledge often occurred through formal cooperation arrangements as well as informal interactions with external actors. This result highlights the transversal role of social integration mechanisms in strengthening the early stages of AC, particularly in international and interorganizational settings.

The findings also indicate that these public servants perform functions analogous to gatekeepers, acting as boundary-spanning actors who connect external environments with internal organizational units. They mediate information flows, translate external practices into organizationally meaningful knowledge, and facilitate the

circulation of strategic insights within the institution. This interpretation is aligned with the gatekeeper concept discussed by Cohen and Levinthal (1990), Daghfous (2004), Graf (2010), and Murray et al. (2011).

Overall, the evidence suggests that AC is concretely manifested in the context of international missions, particularly in the stages related to recognition of value and acquisition of external knowledge. However, although the organization demonstrates strong potential AC, the conversion of individually acquired knowledge into institutional learning depends on subsequent mechanisms of assimilation, transformation, and application. These later-stage constraints are examined in the following section.

4.2 Structural Bottlenecks in Knowledge Transformation and Exploitation

The findings indicate that although international missions function as effective channels for external knowledge acquisition, significant barriers emerge when such knowledge must be internalized, disseminated, and converted into organizational action. The transition from individual learning to institutionalized organizational knowledge remains a challenging process, often interrupted by structural and cognitive limitations related to information processing, coordination, and KM routines.

A critical and often overlooked factor in this dynamic is institutional amnesia. As highlighted by Stark (2018), the erosion of organizational memory caused by leadership changes and personnel turnover undermines policy continuity and weakens accumulated learning. In the absence of robust retention mechanisms, such as institutional memory systems or structured repositories, public organizations face difficulties consolidating past experiences and preserving strategically valuable knowledge over time.

In addition to memory-related constraints, the empirical evidence reveals severe weaknesses in organizational feedback loops. Although public servants assigned abroad routinely submit reports and strategic information, from the experience reported by the interviewees, feedback from central units is frequently fragmented, informal, or absent. According to interviewee 11 "*communication fails both in sending and returning, this ends up generating a disincentive in the production and transfer of knowledge*", leading the servers to seek feedback on their own initiative. This interruption in communication cycles appears associated with structural overload in headquarters units, which often lack sufficient capacity to process, analyze, and respond to the large volume of information generated simultaneously across international representations.

The absence of effective feedback produces an informational vacuum that discourages continued engagement by gatekeepers and weakens strategic alignment between external missions and internal organizational priorities. At the same time, the reception and processing of externally sourced knowledge remain weakly systematized. Reporting practices depend heavily on bureaucratic procedures and individual proactivity, transforming learning into episodic events rather than continuous organizational routines.

Consequently, externally acquired knowledge frequently encounters limited screening capacity, insufficient categorization mechanisms, and low diffusion across internal units. These findings suggest a persistent gap between PACAP and RACAP, in which valuable knowledge is identified and acquired but only partially transformed into institutional learning, innovation, and sustained public value.

4.3 How AI May Strengthen Absorptive Capacity

The empirical findings reveal that the main limitations of AC in the studied public organization are concentrated not in the identification or acquisition of external knowledge, but in its subsequent conversion into institutional learning and organizational action. Gaps in knowledge diffusion, weak feedback cycles, fragmented memory systems, and excessive dependence on individual actors indicate persistent barriers in the transition from PACAP to RACAP. In this context, AI may be understood as a relevant enabling mechanism for strengthening the later stages of absorptive processes.

During assimilation processes, AI may contribute to the organization, classification, and summarization of externally acquired knowledge. One of the recurrent bottlenecks identified in the empirical evidence concerns the large volume of reports, documents, and fragmented information flows that central units often struggle to process. In this regard, natural language processing tools and intelligent repositories may facilitate the codification of dispersed information, transforming heterogeneous inputs into searchable and structured organizational knowledge.

AI may be particularly valuable in addressing institutional amnesia and weak feedback loops, two central constraints identified in the findings. Intelligent memory systems can preserve accumulated experiences, retrieve lessons learned from previous missions, and support continuity despite personnel turnover. Likewise, automated feedback systems may improve communication between central units and staff deployed abroad, helping sustain recursive learning cycles essential to absorptive capacity.

However, the contribution of AI should not be interpreted as purely technological or deterministic. Its effectiveness depends on governance arrangements, data quality, transparency, and human oversight. Most importantly, AI can strengthen absorptive processes only when organizations possess the managerial willingness and organizational routines necessary to interpret and apply digitally supported knowledge.

Therefore, rather than replacing existing human capabilities, AI should be understood as an augmenting capability that supports gatekeepers, strengthens organizational memory, improves knowledge circulation, and facilitates the conversion of externally acquired knowledge into innovation and sustained public value.

The analysis suggests that AI can strengthen the absorptive capacity of public organizations operating in international contexts through three interrelated mechanisms: organizational memory, feedback enhancement, and knowledge dissemination. First, AI systems may support the capture, storage, retrieval, and recombination of knowledge generated through international missions and dispersed networks, helping reduce personnel turnover risks and institutional amnesia. Second, AI may reinforce feedback loops through intelligent filtering, prioritization, and routing of externally acquired knowledge to relevant units, reducing cognitive overload and improving responsiveness in strategic decision-making. Third, AI may expand knowledge dissemination across organizational boundaries through pattern recognition, connection mapping, and collective sense-making from heterogeneous sources.

Figure 1 synthesizes these propositions by illustrating the flow of knowledge from its identification at interorganizational boundaries to its conversion into organizational outcomes. International missions are positioned as gatekeepers, acting as boundary-spanning actors who identify, access, and mediate relevant external knowledge through networks, partnerships, events, and institutional interactions.

This externally sourced knowledge is then processed through AI-enabled mechanisms represented as an intermediate enabling layer. Rather than replacing human actors, AI functions as a support capability that filters information, stores content, facilitates retrieval, identifies patterns, and supports learning processes. In doing so, AI helps organizations manage large volumes of heterogeneous information beyond manual processing capacity. Three reinforcing functions receive particular emphasis in the model:

1. Storage and Retrieval: Representing organizational memory infrastructure, ensuring retention and retrieval of processed knowledge;
2. Feedback: Representing feedback loops. A return arrow connects Feedback directly to gatekeepers abroad, showing that generated intelligence returns to guide new acquisitions and close the learning circuit;
3. Dissemination of Knowledge: Illustrating collective sense-making, where generated intelligence is distributed internally.

Finally, an exit arrow demonstrates that the consolidation of this AI-mediated knowledge governance culminates in incremental innovation and increasing public value, the primary objective of the public organization.

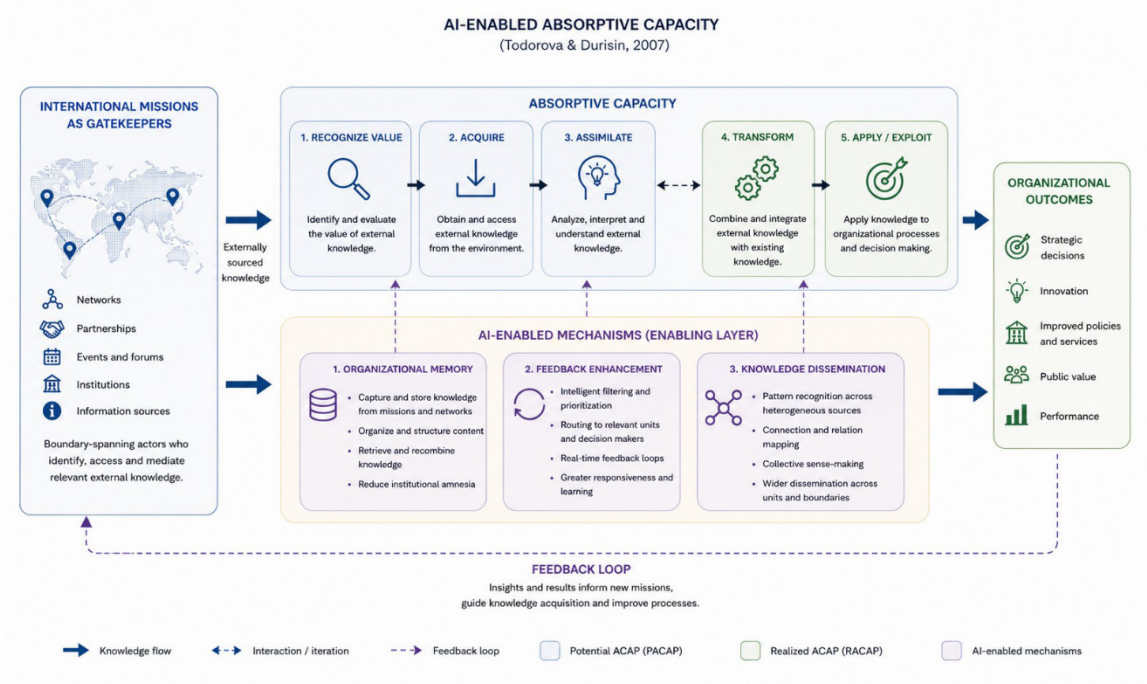


Figure 1: Artificial Intelligence as a vector to strengthen Absorptive Capacity in public context

5 Conclusions

This article achieved its objective by identifying how AC processes unfold in international missions and by proposing how AI may facilitate their different stages in complex public organizations. The results demonstrate that AC is clearly manifested in the context of international missions, particularly in the stages related to recognition of value and acquisition of external knowledge. Public servants assigned abroad operate as gatekeepers, connecting the organization to external knowledge sources. Prior knowledge and technical expertise were shown to be important antecedents supporting these early absorptive processes.

However, the findings also reveal persistent bottlenecks in the later stages of AC. Transformation, and effective application of externally acquired knowledge remain constrained by fragmented routines, weak feedback cycles, insufficient mechanisms for knowledge diffusion, dependence on individual actors, and institutional memory losses associated with personnel turnover. As a result, the organization demonstrates stronger PACAP than RACAP, indicating difficulties in converting external learning into sustained organizational outcomes and public value.

In this context, the study argues that AI may represent an important enabling capability for strengthening absorptive processes. AI tools can support knowledge processing, organizational memory, and decision-making routines, helping public organizations reduce information overload, improve knowledge circulation, and strengthen the transition from PACAP to RACAP. Nevertheless, AI should not be interpreted as a substitute for human judgment, expertise, or relational capabilities. Public servants remain central in interpreting context, exercising discretion, building trust, and translating knowledge into legitimate organizational action. Therefore, the greatest potential of AI lies in augmenting, rather than replacing, human capabilities.

From a theoretical perspective, this article contributes by extending AC research to the context of public-sector international missions and by advancing the dialogue between AC and KM through propositions on how AI may support absorptive processes. From a practical perspective, the findings suggest that managers should focus not only on acquiring external knowledge, but also on developing routines capable of assimilating, transforming, and applying it.

This study presents some limitations. First, it is based on a single-case study within one public organization. Second, the empirical data did not investigate the direct use of AI in the studied context, but treated AI as a conceptual proposition derived from empirical barriers and supporting literature.

Future research may advance this agenda by conducting empirical studies on the actual implementation of AI tools in public organizations and their effects on AC processes. Comparative studies across countries, agencies, and governance contexts may also deepen understanding of how AI interacts with organizational culture, leadership, and institutional constraints. Longitudinal research would be particularly valuable for examining whether AI-supported routines generate sustained learning outcomes over time.

In conclusion, the study suggests that the challenge faced by complex public organizations is not merely to access external knowledge, but to transform it into collective intelligence and public value. In this process, human actors remain indispensable, while AI may serve as a powerful complementary capability for enhancing learning, coordination and innovation.

6 Acknowledgements

We would like to thank the public servers who participated in the interviews for the original study.

7 Ethics declaration

The original research was authorized by the studied organization and approved by the Research Ethics Committee of the Federal University of Santa Catarina, in accordance with National Health Council regulations. Participants were informed about the study objectives, voluntary participation, and confidentiality safeguards through an informed consent form. Data were securely stored and used exclusively for academic purposes. No significant risks or direct benefits were identified, although the study may indirectly contribute to improving knowledge management practices within the organization.

8 AI declaration

Generative Artificial Intelligence tools (ChatGPT and NotebookLM) were used to support language refinement, text structuring, editorial revision, and figure development. All generated outputs were critically reviewed, interpreted, and validated by the authors, who remain fully responsible for the final content, arguments, and conclusions of this study.

References

Ajmal, M. M., ISLAM, Z., ISLAM, A. (2025) Enhancing organizational performance in higher education through knowledge-centered culture and absorptive capacity. *The Learning Organization*, v. 32, n. 5, p. 741-753.

Bardin, L. (2021) *Análise de conteúdo*. Lisboa: Edições 70.

Butler, M.J.R. and Ferlie, E. (2020) 'Developing Absorptive Capacity Theory for Public Service Organizations: Emerging UK Empirical Evidence', *British Journal of Management*, 31(2), pp. 344–364.

Castro, L., Santos-Corrada, M., Flecha-Ortiz, J.A., Lopez, E., Gomez, J. and Aponte, B. (2022) 'Knowledge management and innovative behavior: police reform efforts in Puerto Rico', *Journal of Knowledge Management*, 26(5), pp. 1262-1279.

Cohen, W.M. and Levinthal, D.A. (1989) 'Innovation and learning: the two faces of R&D', *The Economic Journal*, 99(397), pp. 569-596.

Cohen, W. and Levinthal, D. (1990) 'Absorptive capacity: a new perspective on learning and innovation', *Administrative Science Quarterly*, 35(1), pp. 128–152.

Daghfous, A. (2004) 'Absorptive capacity and the implementation of knowledge-intensive best practices', *SAM Advanced Management Journal*, 69(2), pp. 21-27.

Elbashir, M. Z.; Sutton, S. G.; Arnold, V.; Collier, P. A. (2022) Leveraging business intelligence systems to enhance management control and business process performance in the public sector. *Meditari Accountancy Research*, v. 30, n. 4, p. 914-940.

Fagundes, F.M., Tumelero, C. and Fagundes, M.G. (2025) 'Metarrotinas de capacidades absortivas no setor público de saúde', *International Journal of Innovation – IJI*, 13(3), Article e28224.

Françoso, M.S. and Vonortas, N.S. (2023) 'Gatekeepers in regional innovation networks: Evidence from an emerging economy', *The Journal of Technology Transfer*, 48(4), pp. 821-841.

Graf, H. (2011) 'Gatekeepers in regional networks of innovators', *Cambridge Journal of Economics*, 35(1), pp. 173-198.

Jin, Z. and Wang, Y. (2021) 'Resource complementarity, partner differences, and international joint venture performance', International Journal of Strategic Management, 14(2), pp. 255–270.

Joshi, K.D., Chi, L., Datta, A. and Han, S. (2010) 'Changing the competitive landscape: Continuous innovation through IT-enabled knowledge capabilities', *Information Systems Research*, 21(3), pp. 472-495.

Ma, F., Khan, F., Khan, K.U. and XiangYun, S. (2021) 'Investigating the Impact of Information Technology, Absorptive Capacity, and Dynamic Capabilities on Firm Performance: An Empirical Study', *SAGE Open*, 11(3), pp. 1-16.

Mao, H., Liu, S., Zhang, J. and Zhang, Y. (2017) 'Information technology competency and organizational agility: roles of absorptive capacity and information intensity', *Proceedings of the 25th European Conference on Information Systems (ECIS)*, Guimarães, Portugal, pp. 1584-1599.

Mohan, R. (2025) 'Inter-firm imitation of artificial intelligence: Towards innovation and competitive edge in business', *Journal of Business Innovation*, 12(4), pp. 700–715.

Mori, S.S.O., Santos, C.D. and Ferraz, I.N. (2025) 'Efeitos dos mecanismos de governança na capacidade absortiva de projetos de colaboração internacional: um estudo na organização Embrapa', *Revista Brasileira de Inovação*, 24, e025012.

Murray, K., Roux, D.J., Nel, J.L., Driver, A. and Freimund, W. (2011) 'Absorptive capacity as a guiding concept for effective public sector management and conservation of freshwater ecosystems', *Environmental Management*, 47(6), pp. 917-925.

Pinho, P.H.S., Sousa-Filho, J.M., Lessa, B.S. and Costa, J.S. (2024) 'Absorptive Capacity, Entrepreneurial Orientation and organizational performance in State-Owned companies', *International Journal of Public Sector Management*, 37(5), pp. 677–691.

Stark, A. (2018) 'Explaining institutional amnesia in government', *Governance*, 32(1), pp. 143-158.

Todorova, G. and Durisin, B. (2007) 'Absorptive capacity: Valuing a reconceptualization', *Academy of Management Review*, 32, pp. 774–786.

Van der Graaf, P., Cheetham, M., Redgate, S., Humble, C. and Adamson, A. (2021) 'Co-production in local government: process, codification and capacity building of new knowledge in collective reflection spaces. Workshop findings from a UK mixed methods study', *Health Research Policy and Systems*, 19, Article 12.

Waheed, J. (2025) Linking absorptive capacity and project performance in environmental uncertainty: A perspective on implementation arrangements. *Review of Policy Research*, v. 42, p. 252-275.

Zahra, S.A. and George, G. (2002) 'Absorptive Capacity: A review, reconceptualization, and extension', Academy of Management Review, 27, pp. 185–203.

Zhou, L., Chen, L., Han, Y. (2021) 'Data stickiness" in interagency government data sharing: a case study, *Journal of Documentation*, 77(6), pp. 1286-1303.

Influence of Artificial Intelligence (AI) on Organizational Knowledge Creation (OKC) in Workplace Meetings

Shuyun Zhao[1], Krishna Venkitachalam[1,2]
[1] Estonian Business School, Tallinn, Estonia
[2] Ajman University, UAE

Abstract

Workplace meetings are important to organizations' daily operations, and knowledge creation is a valuable outcome of these meetings. Despite the increasing use of AI in meetings, its influence on OKC remains unexamined. This study uses the SECI model as the theoretical framework and aims to explore how AI influences the OKC process before, during, and after workplace meetings.

This study adopted a qualitative interpretive approach to explore and understand a new and complex phenomenon. Eight semi-structured interviews were conducted between March and April 2026 with employees from 3 countries and across different industries who regularly use AI in work-related meetings. Data were analyzed using the thematic analysis method with deductive-inductive coding.

The findings show that AI encourages people to engage in social interaction. In externalization, AI enables the articulation process via several mechanisms. AI enhances combination by connecting large amounts of existing codified knowledge from multiple sources. In the last mode, AI empowers individuals in internalizing knowledge. A conceptual framework is proposed that maps AI's influence in the SECI model and indicates the primary work meeting phases.

The study's qualitative nature and narrow scope restrict how broadly its findings can be statistically generalized. Second, it focuses solely on the SECI model as the analytical lens. Third, the rapid evolution of AI means the findings are time-bound. Future research can focus on AI-facilitated meetings across a wider array of organizations, apply diverse theoretical models to OKC, and scrutinize the dynamic effects of AI on this phenomenon.

The study extends the SECI model by including AI's influence in the four knowledge conversion modes in workplace meetings. Essentially, it sheds light on how AI in meetings reshapes the way organizations create knowledge, possibly enabling more effective AI application and use.

Keywords: artificial intelligence, organizational knowledge creation, SECI, workplace meetings

Paper type: Academic Research Paper

1 Introduction

Workplace meetings refer to "purposeful work-related interactions occurring between at least two individuals that have more structure than a simple chat but less structure than a lecture" (Rogelberg et al., 2025, p. 416). Researchers have identified various types of meetings. They can be differentiated by whether they are scheduled or unscheduled, or by different purposes such as discussion of ongoing projects, routine business status updates, quality, policy, and compliance. With advances in technology, virtual and hybrid meetings are becoming increasingly popular beyond traditional in-person meetings (Abidi et al., 2023). Meetings are a core component of how organizations function, irrespective of their cultural background, industry sector, or scale. They can serve as a forum for idea generation, debate and discussion, and creative thinking, leading to action plans for the next steps (Lawson et al., 2009), with the creation of new knowledge as a valuable outcome. The adoption of AI tools for meeting mediation has grown rapidly, encompassing those for meeting preparation, real-time transcription, summary generation, and support for follow-up activities. The growing use of AI is reshaping how knowledge is created in workplace meetings.

Organizational studies suggest that AI will work in synergy with human workers (Jarrahi et al., 2023), especially with the development of generative AI (Alavi et al., 2024). AI supports all four modes of the SECI model to varying degrees. For example, AI can facilitate socialization by removing physical constraints among remote teams or by connecting people with shared interests (Böhm and Durst, 2026). In externalization, people's tacit knowledge is converted to explicit knowledge via continuous dialogue with ChatGPT (Sumbal and Amber, 2025). In combination, AI can integrate fragmented resources (Almufadda and Almezeini, 2022), analyze large data sets, and uncover new patterns (Faulconbridge et al., 2024). Lastly, AI supports internalization by providing personalized assistance, recommendations, and tailored explanations and examples (Sumbal and Amber, 2025). Meetings exemplify the four modes: socialization through direct interaction, externalization by articulating thoughts, combination by integrating different knowledge bases, and internalization by applying new concepts to one's work. However, AI's role in knowledge creation in meetings is significant and warrants further

examination. Therefore, this study aims to explore the following research question: How does AI influence the organizational knowledge creation process before, during, and after workplace meetings?

The findings reveal that AI has an uneven influence on organizational knowledge creation in workplace meetings across the four SECI modes and in different meeting phases: it encourages socialization, enables articulation, enhances combination by linking various knowledge sources, and empowers individuals in internalization. This paper extends the SECI model by incorporating the influence of AI in each mode and indicating the primary meeting phase of its occurrence. It also highlights the importance of human-AI collaboration in creating new knowledge. This paper is organized as follows. The following section reviews the literature on OKC and its relevance to AI-facilitated meetings. The methodology part explains how data were collected through semi-structured interviews and analyzed using the thematic approach. The findings section is organized around four themes as per the SECI model, followed by a conceptual framework. The discussion and conclusions are presented at the end.

2 Literature Review

2.1 Organizational knowledge creation

OKC is commonly considered the first step for managing organizational knowledge and plays an important role in fostering innovation (Bettiol et al., 2023). By definition, it is "the process of making available and amplifying knowledge created by individuals as well as crystallizing and connecting it with an organization's knowledge system" (Nonaka et al., 2006, p. 1179). The process involves the conversion between tacit and explicit knowledge – tacit knowledge is deeply embedded in one's experiences, intuition, and expertise and is hard to articulate, while explicit knowledge is codified or articulated in a formal manner (Venkitachalam and Schiuma, 2022).

The SECI model depicts four knowledge conversion modes: socialization, externalization, combination, and internalization (Nonaka and Takeuchi, 1995). In socialization, through shared experiences, interactions, and observations, tacit knowledge is transferred among individuals; externalization is the process by which tacit knowledge is codified or articulated into explicit form; in combination, existing explicit knowledge is reorganized and synthesized into more systematic forms; internalization is the process for individuals to learn from the reorganized explicit knowledge and absorb it as their own tacit knowledge (e.g. Nonaka and Toyama, 2003, Lievre and Tang, 2015, Canonico et al., 2020). These processes remain important today as AI shapes how knowledge is created in organizations. For example, Zhang et al. (2025) expanded the model incorporating AI, using interpretation Ba and decision-making Ba to replace the original ones in externalization and combination. He et al. (2025) proposed that generative AI serve as a Ba by facilitating the interaction among employees. Workplace meetings provide a shared space for knowledge creation, particularly in virtual or hybrid formats, and offer an important lens for examining how AI influences the OKC process.

2.2 Workplace meetings and AI mediation - Context for OKC

Workplace meetings are key to organizational operations, which are shaped at the individual, group/team, and organizational levels. They also provide a context in which the organizational knowledge-creation process can be directly observed. In a two-year study of 43 team meetings, Hubers et al. (2016) demonstrated that meetings enabled all four modes in the SECI model, while identifying the iterative nature of knowledge creation within meetings and the cyclical nature across meetings. Each knowledge conversion mode has its distinctive characteristics in meetings, as the remainder of this section demonstrates.

First, meetings serve as an important socialization mechanism that connects individuals and promotes close interactions, thereby encouraging the exchange of knowledge (Lawson et al., 2009). Through repeated interactions, meetings help participants build trust, respect, shared norms, and mutual understandings - all of which are preconditions for knowledge creation. Socialization requires active maintenance – in effective teams, this is evident in pre-meeting interactions, social activities during meetings, and follow-up mechanisms. In remote teams, however, the extent of socialization is greatly reduced compared to co-located teams for various reasons. As a result, organizations rely on collaborative technologies such as teleconferencing, videoconferencing, and electronic meeting systems to increase socialization (Oshri et al., 2007).

Next, meetings are places where tacit knowledge is expressed and heard by peers (Nonaka and Konno, 1998). This is done via three complementary mechanisms: verbal, visual, and written. First, meetings are fundamentally constituted by talk (Asmuß and Svennevig, 2009). Second, visual artifacts, such as PowerPoint presentations (Kaplan, 2011) or similar tools, can be used for articulating one's thinking. Third, meetings typically involve

various written documents, such as a written agenda outlining the topics to be discussed before the meeting and minutes after the meeting (Gutmann, 2016). However, the externalization of tacit knowledge in meetings can be constrained, for example, not everyone's voice can be heard (Bohacova and Heide, 2024).

Third, combination occurs in meetings by organizing and integrating explicit knowledge from different participants or sources into more synthesized forms (Nonaka and Takeuchi, 1995). Various explicit inputs are brought to meetings, such as handouts, graphs, and reports (Henkel, 2007). In turn, various new explicit artifacts, such as minutes, notes, and action lists, are produced after meetings (Gutmann, 2016), which can be fed into subsequent meetings (Jarzabkowski and Seidl, 2008).

Fourth, internalization is less evident during meetings but more evident after meetings, when participants need to apply explicit knowledge to their work practices. Even though internalization can be less dominant during meetings, team members were observed to briefly reflect on the project progress or their own learning process (Hubers et al., 2016). This mode involves embodying explicit knowledge into action (Nonaka and Konno, 1998), for example, by developing action plans for employees to follow after meetings (Kauffeld and Lehmann-Willenbrock, 2012). In addition, as collaborative work passes through a series of successive meetings (Jarzabkowski and Seidl, 2008), the internalized knowledge is carried back to discussions, which iterates the knowledge creation process (Nonaka and Takeuchi, 1995).

AI is increasingly integrated into meetings, reshaping how knowledge is created before, during, and after meetings. Meeting-specific AI assistants, such as those in Zoom, Google Meet, and Microsoft Teams, provide live transcription, translation, and summarization (Suduc and Bizoi, 2022), while generative AI tools like ChatGPT can enhance the four modes of the SECI model (Sumbal and Amber, 2025), supporting both meeting preparation and post-meeting knowledge internalization. However, using AI has its shortcomings, especially around the nuances of knowledge. Minute-taking, for example, is not just about recording notes but also about prioritizing topics contextually and building a common understanding among. Transcription accuracy is another common issue, which may be affected by the use of different languages (Arakawa et al., 2025). While prior literature illustrates both AI's the advantages and the shortcomings, its influence on organizational knowledge creation remains underexplored.

3 Methodology

This study examines how employees experience and interpret AI-mediated knowledge creation in workplace meetings. Because the phenomenon of adopting AI in meetings is recent and highly context-dependent, this study adopts a qualitative interpretive approach (Myers, 2025). We conducted semi-structured interviews with employees with regular experience using AI in work-related meetings. Qualitative interviews allow the flexibility to explore themes further and capture unanticipated insights. We employed a convenience sampling method: participants were recruited through the researchers' personal networks, followed by snowball sampling. In total, we conducted 8 interviews between March and April 2026 with participants in diverse roles across organizations in three countries and across different industries. Table 1 illustrates the background of participants.

Table 1: Profile of participants

Interview	Participant's Role	Industry	Organization Country
1	Senior technical expert	Autonomous driving	China
2	Software engineer	Healthcare	China
3	Legal counsel	Entertainment	China
4	Chief governance officer	Fintech	Estonia
5	Staff software engineer	Technology	Estonia
6	Legal manager	Cosmetics	China
7	Strategic business architect	Financial services	Estonia
8	Brand partnerships manager	Social media	UK

The interviews were conducted via videoconference platforms (Zoom or iFLYREC) in either Chinese or English, according to the participant's preference, and lasted 30-80 minutes, averaging 59 minutes. The conversations followed a semi-structured interview guide built around a set of open-ended questions and grounded in the SECI model, allowing the interviewer to probe and ask follow-ups to clarify meaning and pursue unanticipated themes. The interviews were audio-recorded with participants' consent and fully transcribed. For the interviews conducted in Chinese, the original language was used for coding and analysis, to preserve participants' own expressions and avoid interpretive loss (Abfalter et al., 2021). The Chinese transcripts were translated into English with AI assistance and reviewed word by word to ensure faithful representation of participants' meaning.

We used thematic analysis (Braun and Clarke, 2006) to analyze the collected data. A hybrid coding approach (Fereday and Muir-Cochrane, 2006) was used. In line with the interview guide, the initial codebook includes a set of a priori codes derived from three dimensions: the SECI framework, the temporal phases of meetings, and participants' evaluations of AI's effect on organizational knowledge creation. Additional codes were generated inductively as new patterns emerged. We iteratively refined the codebook by applying the newly surfaced codes. This process is done via the tool ATLAS.ti.

4 Findings

The findings show that the four SECI modes unfold unevenly across different meeting phases, while AI is also used to varying degrees before, during, or after meetings. Accordingly, AI's influence on organizational knowledge creation in different meeting phases is unevenly distributed. In total, four key themes emerged from the data, one for each SECI mode.

4.1 Theme 1: Socialization is AI-encouraged

On the one hand, AI has a limited influence on socialization, supported by strong participant opinions such as *"whether AI is involved doesn't make much difference" (Interview 1)*, *"I really don't think this has much to do with AI" (Interview 5)*, and *"it will not affect interaction with others at all" (Interview 3)*. The constraint lies in the nature of tacit knowledge, which is deeply embedded in human minds. Examples include detecting whether a candidate is using AI to answer during interview meetings, or deciding which product development demand is the most urgent.

"… I'm only recording the transcription, that's it. But what I'm doing is looking at the body language of the candidate … they cannot hide their eyes … when they answer, they always look in a very concrete direction, and then I automatically decline." (Interview 4)

"… The first human judgment factor is the urgency of the requirement. Business teams usually say everything is urgent… You will have to make your own judgment … or more broadly, your work experience." (Interview 2)

On the other hand, participants acknowledge the indirect supporting role of AI in specific aspects of socialization. The first is that AI can sometimes help participants read between the lines of obscure communication, whether in internal meetings with managers, colleagues, or external meetings with customers (Interview 7, 8).

"… when British people make decisions, especially those involving risk, they may not explicitly state it … AI will then interpret the manager's words … it may summarize that your manager currently prefers all three options together, or that all three options are feasible. So, my suggestion would be choosing the one you personally prefer and then confirming it again with your manager." (Interview 7)

Second, AI reallocates meeting participants' attention from notetaking to the people in the room, also allowing the speaker to focus on expressing their ideas without interruption.

"… You may need to interrupt the process to take notes. But if you use AI, this problem basically does not exist anymore, because I know AI is helping me record everything." (Interview 2)

Third, AI subtly smooths communication by helping people understand each other by removing the barrier of accent or listening ability (Interview 5), removing background noise or showing captions during internet issues (Interview 6), and providing real-time translation (Interview 4, 8). It can also give people the freedom to learn each other's fields before or after meetings, reducing the need for explanations during the meetings.

"… With AI's help, one can learn about that field beforehand, so the meeting is more efficient." (Interview 1)

"… For example, some things don't necessarily need to be explained in full during a meeting. After the meeting … they can just use AI to summarize it. This may save some time in the socialization process." (Interview 5)

In summary, socialization – the conversion of tacit knowledge among individuals through shared experience or interactions – is the mode that has been least transformed by AI in the meeting context. Yet AI can indirectly support this mode by encouraging some interaction.

4.2 Theme 2: Externalization is AI-enabled

Documentation is key to externalization, as it efficiently turns tacit knowledge into written texts that can be stored for future use. This process generates different types of explicit outputs, including original audio or video recordings, transcripts, or a minute or summary.

"… an important way of turning tacit knowledge into explicit knowledge is documentation … AI saves a lot of effort, particularly in that process of textualization." (Interview 1)

"…It keeps the original audio … provides the full meeting transcript translated from the original recording … presents a summarized version as well. So, it gives you all three versions." (Interview 2)

"… First help is transcribing it, quickly and qualitatively. Second is restructuring and organizing it in a reusable form." (Interview 4)

In terms of quality, AI is considered a better notetaker than ordinary people. It not only generates easy-to-read summaries quickly but also filters out irrelevant content. From the users' points of view, they do not need to worry about grammar or flow.

"… AI is very powerful when it comes to note-taking or summarizing. The notes generated by AI are usually quite readable. In contrast, when you take notes yourself … the result is not very easy to read." (Interview 2)

"… when you say something to ChatGPT, you don't have to pay attention to grammar, and what you say might be a little incoherent, but you know the AI understands you." (Interview 5)

"… They will use AI to help make the document flow more smoothly and logically. In addition, non-native English speakers may use AI to help correct grammar mistakes … Although the person is still the main creator, AI helps package the content in a more refined and polished way." (Interview 8)

However, the notes generated by AI are not without flaws - one thing frequently mentioned by several participants is the loss of meaning, importance, or context in the summaries it produces. One participant illustrated the situation with a backward-installed camera case, a serious error that needs urgent correction. The top priority makes it unnecessary to explain anything else; by merely mentioning this issue, everyone knows what to do immediately after the meeting. However, for AI, this is a simple fact: AI may or may not include it in the summaries (Interview 1). In a similar vein, AI misses the priorities of tasks in developing demands, which are usually decided by project managers (Interview 2). In the case where there is a lack of company-specific context, such as internal adjustments or customizations, it is difficult for AI to accurately externalize them (Interview 5). To improve the quality, users need to provide contextual content to AI.

"… we would write a few A4 pages of content, copy and paste it into AI first, and set that as our background information … We had a saying at the time, … 'Learn to prompt, prompt to learn.' … meaning you need to explain what your context is." (Interview 7)

"… accuracy is related to the glossary. That means names, company-specific terms, personal names, or other specific terminology. We manage the glossary separately. We have taught AI to use certain terms in specific ways. Every time I find a mismatch … I add that term to a glossary that AI can access." (Interview 4)

In summary, externalization – the process that converts individuals' tacit knowledge into explicit form – is the mode that has been strongly transformed by AI in the meeting context. It claims that AI strongly enables externalization via several mechanisms, such as generating transcriptions and minutes, with output quality generally satisfactory though sometimes requiring human editing. However, because of the nature of tacit knowledge, AI can lose meaning, importance, or context.

4.3 Theme 3: Combination is AI-enhanced

AI enhances the combination process by effectively and efficiently synthesizing across various sources. Before AI, one participant's workflow using search engines was: opening a webpage, checking its relevance, opening more webpages to identify better sources, or making a summary. Now the tasks are done by AI - although the underlying process is somewhat similar, it becomes far more efficient (Interview 2). This efficiency is also evident within organizations. In the past, internal portals such as wikis or SharePoint were used to maintain a centralized knowledge base, which required continuous management and special efforts. AI can now integrate knowledge from more scattered internal sources, including memos, policies, procedures, and standards. In this way, AI reduces manual effort while also improving the richness of knowledge (Interview 4). Another participant also stresses AI's strong capability to integrate knowledge from the company's internal knowledge base (Interview 5).

However, participants stress that AI's influence on knowledge combination occurs primarily in the pre- or post-meeting phases, not during the meeting itself, because meetings often focus on a specific topic or subtask rather than on cross-field synthesis. In practice, the knowledge generated from meetings may be transformed into other organizational artifacts.

"… Any meeting should have a clear topic. It should focus on a small theme, a specific direction, a sub-task, … there is usually little cross-disciplinary or cross-topic discussion … After the summary is done and the meeting ends, the key question is what to do with the documents and knowledge that have been produced and accumulated from the meeting." (Interview 1)

"…Usually, the knowledge we gain from meetings ends up elsewhere. It may be turned into development tasks, business requirements, or similar items. Then it is decomposed and stored elsewhere. So later, I am more likely to review those sources, rather than the meeting outputs directly." (Interview 4)

In summary, combination - the synthesis of existing explicit knowledge into a more organized form - is enhanced by AI, though constrained by the meeting context. AI helps integrate and synthesize explicit knowledge. However, because of the scope of meetings, the use of AI in this mode is not significant.

4.4 Theme 4: Internalization is AI-empowered

Participants express the idea that AI empowers the internalization process, which mostly occurs after meetings. One participant gave the example of after-training meetings, describing AI's role as a "tutor" or "study buddy" that helps your learning using existing material (Interview 1). The other participant has stopped using traditional search engines and is using AI instead to help master new knowledge. By comparing the results generated by different AI agents and finding the most suitable results, the participant can be further empowered (Interview 2).

"… before AI, you really had to spend a lot of time reading code, talking to people, or reading internal documents before you could become a very experienced employee. But now, … the threshold has become much lower. You can just ask AI … it can basically act as a proxy for an expert." (Interview 5).

However, despite the positive attitudes towards the benefits of AI in accelerating internalization, there is one negative example worth noting. "I became lazy", as stated by one participant, expressing the worry of the weakening influence of AI on an individual's motivation to acquire new knowledge.

"… Without AI, I would actively try to learn a lot of new knowledge … As a result, my understanding of that issue would become more comprehensive … I can clearly feel that this is an obvious difference: people really have become lazier, because knowledge has become too easy to obtain." (Interview 7)

In summary, people are empowered by AI in the process of internalization – the process that involves the absorption of explicit knowledge into one's tacit knowledge. It claims that AI serves as a companion in the internalization mode; however, the negative influence of AI is that the overdependency on AI weakens one's own ability and motivation to learn.

5 Toward a Conceptual Framework - Discussion and Conclusion

Based on the thematic findings, we propose a conceptual framework that depicts AI's influence on knowledge creation in workplace meetings (see Figure 1). It integrates two aspects: first, a mapping of AI's influence across the SECI modes; second, a temporal dimension indicates the primary meeting phase(s) during which AI influences knowledge creation.

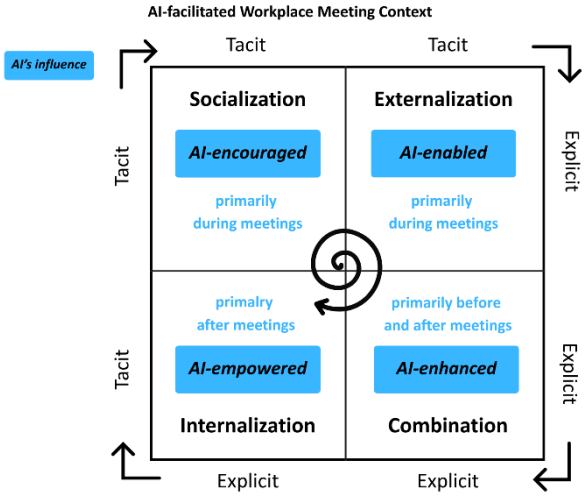


Figure 1: AI's influence in SECI-based framework of organizational knowledge creation in workplace meetings

The diverse impact of AI across the SECI modes in workplace meetings offers a profound understanding: humans and AI can partner in various ways to generate novel knowledge. This bears a strong resemblance to the contrast between tacit and explicit knowledge. AI handles codified knowledge well, but humans remain central for tacit knowledge, as socialization shows. "Learn to prompt, prompt to learn" – a slogan coined by one participant in the company to encourage people to better collaborate with AI (Interview 7) – captures the nature

of this collaboration. Currently, AI might struggle with the nuances of specific organizational contexts, which can affect its results. Metaphors from participants suggest diverse human-AI collaboration models: an assistant, a manual worker, or J.A.R.V.I.S. (the AI assistant from the Iron Man movies); second, a tutor or study buddy; and the last, a knowledge expert, a wise man.

By examining these patterns, the study deepens our understanding of AI's impact on organizational knowledge creation in workplace meetings. The proposed framework demonstrates AI's influence on knowledge creation within each SECI mode. Relative to other SECI-based frameworks, including those re-examining Ba with AI (Zhang et al., 2025, He et al., 2025), and GRAI (Generative, Receptive Artificial Intelligence) as an evolutionary step beyond SECI (Böhm and Durst, 2026). The SECI framework's fundamental conversion modes remain unchanged by our approach, but we broaden its application by adding AI's impact on workplace meetings. For example, the study's practical implications highlight the importance of employees understanding how to communicate with and instruct AI for improved knowledge creation.

This study has a few acknowledged limitations. The study's qualitative nature and narrow scope limit the statistical applicability of its findings. It exclusively uses the SECI model to analyze how organizations create knowledge, whereas other theories can be considered for future research. Third, AI technology is advancing rapidly, meaning the results are not current and might overlook AI's expanding influence on how organizations generate knowledge. Further research should investigate AI-assisted meetings across diverse organizations or contrast findings between various organizational types. Alternative theories or frameworks for how organizations create knowledge can also provide analytical viewpoints. Given AI's ongoing development, we should examine its increasing role in workplace knowledge creation during meetings and beyond.

6 Ethics Declaration

The research did not need ethical clearance.

7 AI Declaration

The initial transcripts were generated by the videoconferencing platforms (Zoom and iFLYREC) with their built-in AI transcription tools. Claude Opus 4.7 was used to clean the transcripts by removing filler words and marking low-confidence sections, and to translate Chinese transcripts to English. The transcripts were reviewed verbatim against the original recordings, and the translations were reviewed word by word.

References

Abfalter, D., Mueller-Seeger, J. and Raich, M. (2021) "Translation Decisions in Qualitative Research: A Systematic Framework". *International Journal of Social Research Methodology*, Vol. 24**,** No. 4, pp 469-486.

Abidi, O., Dženopoljac, V. and Safi, M. (2023) "Online Meeting Tools, Tacit Knowledge Sharing and Entrepreneurial Behaviours among Knowledge Workers during Covid-19". *Knowledge Management Research & Practice*, Vol. 21**,** No. 6, pp 1137-1149.

Alavi, M., Leidner, D. and Mousavi, R. (2024) "A Knowledge Management Perspective of Generative Artificial Intelligence". *Journal of the Association for Information Systems*, Vol. 25**,** No. 1, pp 1-12.

Almufadda, G. and Almezeini, N. A. (2022) "Artificial Intelligence Applications in the Auditing Profession: A Literature Review". *Journal of Emerging Technologies in Accounting*, Vol. 19**,** No. 2, pp 29-42.

Arakawa, R., Yakura, H., Akuzawa, K. and Kubo, S. Ai for Meeting Minutes: Promises and Challenges in Designing Human-Ai Collaboration on a Production Saas Platform. Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems, 2025. 1-6.

Asmuß, B. and Svennevig, J. (2009) "Meeting Talk: An Introduction". *Journal of Business Communication*, Vol. 46**,** No. 1, pp 3-22.

Bettiol, M., Capestro, M., Di Maria, E. and Grandinetti, R. (2023) "Leveraging on Intra- and Inter-Organizational Collaboration in Industry 4.0 Adoption for Knowledge Creation and Innovation". *European Journal of Innovation Management*, Vol. 26**,** No. 7, pp 328-352.

Bohacova, K. and Heide, M. (2024) "The Silent Potential: Coworkers' Voicing in Workplace Meetings". *Corporate Communications: An International Journal*, Vol. 29**,** No. 7, pp 130-147.

Böhm, K. and Durst, S. (2026) "Knowledge Management in the Age of Generative Artificial Intelligence–from Seci to Grai". *VINE Journal of Information and Knowledge Management Systems*, Vol. 56**,** No. 1, pp 106-121.

Braun, V. and Clarke, V. (2006) "Using Thematic Analysis in Psychology". *Qualitative research in psychology*, Vol. 3**,** No. 2, pp 77-101.

Canonico, P., De Nito, E., Esposito, V., Iacono, M. P. and Consiglio, S. (2020) "Knowledge Creation in the Automotive Industry: Analysing Obeya-Oriented Practices Using the Seci Model". *Journal of Business Research*, Vol. 112**,** No., pp 450-457.

Faulconbridge, J. R., Sarwar, A. and Spring, M. (2024) "Accommodating Machine Learning Algorithms in Professional Service Firms". *Organization Studies*, Vol. 45**,** No. 7, pp 1009-1037.

Fereday, J. and Muir-Cochrane, E. (2006) "Demonstrating Rigor Using Thematic Analysis: A Hybrid Approach of Inductive and Deductive Coding and Theme Development". *International journal of qualitative methods*, Vol. 5**,** No. 1, pp 80-92.

Gutmann, J. (2016) *Taking Minutes of Meetings*, London, Kogan Page Publishers.

He, X., Antonczak, L. and Burger-Helmchen, T. (2025) "Revisiting Ba: How Generative Ai Transforms Knowledge Creation". *Journal of Knowledge Management*, Vol. 30**,** No. 2, pp 846–869.

Henkel, S. L. (2007) *Successful Meetings: How to Plan, Prepare, and Execute Top-Notch Business Meetings*, Atlantic Publishing Company.

Hubers, M. D., Poortman, C. L., Schildkamp, K., Pieters, J. M. and Handelzalts, A. (2016) "Opening the Black Box: Knowledge Creation in Data Teams". *Journal of Professional Capital and Community*, Vol. 1**,** No. 1, pp 41-68.

Jarrahi, M. H., Askay, D., Eshraghi, A. and Smith, P. (2023) "Artificial Intelligence and Knowledge Management: A Partnership between Human and Ai". *Business Horizons*, Vol. 66**,** No. 1, pp 87-99.

Jarzabkowski, P. and Seidl, D. (2008) "The Role of Meetings in the Social Practice of Strategy". *Organization studies*, Vol. 29**,** No. 11, pp 1391-1426.

Kaplan, S. (2011) "Strategy and Powerpoint: An Inquiry into the Epistemic Culture and Machinery of Strategy Making". *Organization Science*, Vol. 22**,** No. 2, pp 320-346.

Kauffeld, S. and Lehmann-Willenbrock, N. (2012) "Meetings Matter: Effects of Team Meetings on Team and Organizational Success". *Small Group Research*, Vol. 43**,** No. 2, pp 130-158.

Lawson, B., Petersen, K. J., Cousins, P. D. and Handfield, R. B. (2009) "Knowledge Sharing in Interorganizational Product Development Teams: The Effect of Formal and Informal Socialization Mechanisms". *Journal of Product Innovation Management*, Vol. 26**,** No. 2, pp 156-172.

Lievre, P. and Tang, J. (2015) "Seci and Inter-Organizational and Intercultural Knowledge Transfer: A Case-Study of Controversies around a Project of Co-Operation between France and China in the Health Sector". *Journal of Knowledge Management*, Vol. 19**,** No. 5, pp 1069-1086.

Myers, M. D. (2025) *Qualitative Research in Business and Management*, London, SAGE Publications Ltd.

Nonaka, I. and Konno, N. (1998) "The Concept of "Ba": Building a Foundation for Knowledge Creation". *California Management Review*, Vol. 40**,** No. 3, pp 40-54.

Nonaka, I. and Takeuchi, H. (1995) *The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation*, New York, Oxford University Press.

Nonaka, I. and Toyama, R. (2003) "The Knowledge-Creating Theory Revisited: Knowledge Creation as a Synthesizing Process". *Knowledge Management Research & Practice*, Vol. 1**,** No. 1, pp 2-10.

Nonaka, I., Von Krogh, G. and Voelpel, S. (2006) "Organizational Knowledge Creation Theory: Evolutionary Paths and Future Advances". *Organization Studies*, Vol. 27**,** No. 8, pp 1179-1208.

Oshri, I., Kotlarsky, J. and Willcocks, L. P. (2007) "Global Software Development: Exploring Socialization and Face-to-Face Meetings in Distributed Strategic Projects". *The Journal of Strategic Information Systems*, Vol. 16**,** No. 1, pp 25-49.

Rogelberg, S., Kreamer, L. M. and Gray, J. (2025) "Thirty Years of Meeting Science: Lessons Learned and the Road Ahead". *Annual Review of Organizational Psychology and Organizational Behavior*, Vol. 13**,** No., pp 415-442.

Suduc, A. M. and Bizoi, M. (2022) "Ai Shapes the Future of Web Conferencing Platforms". *Procedia Computer Science*, Vol. 214**,** No., pp 288-294.

Sumbal, M. S. and Amber, Q. (2025) "Chatgpt: A Game Changer for Knowledge Management in Organizations". *Kybernetes*, Vol. 54**,** No. 6, pp 3217-3237.

Venkitachalam, K. and Schiuma, G. (2022) "Strategic Knowledge Management (Skm) in the Digital Age–Insights and Possible Research Directions". *Journal of Strategy and Management*, Vol. 15**,** No. 2, pp 169-174.

Zhang, W., Zhang, W., Daim, T. and Yalçın, H. (2025) "Ai Challenges Conventional Knowledge Management: Light the Way for Reframing Seci Model and Ba Theory". *Journal of Knowledge Management*, Vol. 29**,** No. 5, pp 1618-1654.

Beyond Marketplace Reviews: Transformer-Based Sentiment Analysis of YouTube Influencer Ecosystems

Laura Nucera, Davide D'Emiliano, Andrea De Mauro
LUISS University, Rome, Italy

Abstract

Online reviews and influencer-centred social media conversations jointly shape product understanding in digital markets, yet they are still often studied as separate streams. This gap is especially relevant in the beauty sector, where product evaluation depends on subjective experience, perceived efficacy, tolerability, routine fit, and trust in peer-generated accounts. This study develops a cross-platform comparison of Amazon reviews and YouTube comments related to three active skincare products from the same brand, to examine whether the emotional profiles emerging from marketplace reviews differ from those found in influencer-centred comment spaces, and whether YouTube comments reveal needs, doubts, and advice-seeking signals that remain less visible in marketplace environments. The analysis is based on a replicable pipeline that combines data collection, filtering, text preparation, and transformer-based emotion detection using RoBERTa GoEmotions. The findings show that Amazon and YouTube do not produce equivalent emotional profiles. Across the three products, Amazon tends to concentrate emotional expression around admiration and approval, whereas YouTube comments display a broader and more fragmented emotional repertoire. The paper contributes theoretically by integrating review research and influencer ecosystems within a common comparative perspective, methodologically by proposing a transparent cross-platform workflow, and managerially by showing how YouTube-centred social listening can support the detection of consumer concerns that remain underrepresented in marketplace reviews.

Keywords: business model; digital platforms; human–AI interface; intelligent knowledge; data governance.

Paper type: Academic Research paper

1 Introduction

In contemporary digital ecosystems, online reviews are central to consumer decision-making. Despite this centrality, much of the literature continues to favour sentiment analysis of text-based reviews published on brand-led environments like Amazon, while paying less attention to evaluative content produced in influencer-led communication spaces, such as YouTube. This gap is particularly relevant in the beauty sector, where brands operate within multi-platform ecosystems in which company-sanctioned spaces coexist with more spontaneous, socially mediated discursive environments. In this scenario, YouTube's relevance to the customer journey is supported by market evidence and literature, which suggests that conversations around video reviews act as an 'extended focus group,' producing rich insights into consumer needs, preferences, and evaluations (Kabel et al., 2024). From these premises, the research gap that motivates this study is twofold. First, the literature still tends to treat marketplace reviews and conversations developed in influencer spaces separately, without building a systematic comparative framework between sources that, from the consumer's perspective, jointly contribute to the formation of product perceptions. Second, much research continues to operate through aggregate categories of positive or negative polarity, which are not always sufficient to capture finer differences in emotional profiles. Emotions with the same valence can produce different effects, and emotional language is often contextual and difficult to reduce to simplified classifications (Mustak et al., 2024). For this reason, cross-platform comparison requires an analytical framework capable of preserving differences across socio-technical contexts while simultaneously making different forms of evaluative expression comparable. From this perspective, transformer models such as BERT and RoBERTa are particularly relevant for analysing sentiment and emotions, although they require explicit attention to the limitations and biases of available benchmarks (Aka Uymaz & Kumova Metin, 2022; Demszky et al., 2020). In light of this, the study develops a cross-platform comparative analysis of online opinions on beauty products. It compares texts from company-sanctioned platforms and influencer spaces. The goal is to understand how each environment contributes to the development of emotional sentiment profiles and which emotional dimensions differ across channels. The expected contribution is threefold. Theoretically, the study links eWOM, online reviews, and research on influencer communication. Methodologically, it offers a comparative logic for analysing diverse texts. Empirically and managerially, a multi-source sentiment reading helps identify emerging themes, perceived issues, and weak signals.

The research questions follow from this framework and guide comparison between official brand communications and influencer-led social conversations:

- *RQ1: How do emotional profiles differ between marketplace review environments and influencer-led social conversations in beauty product discourse?*
- *RQ2: What consumer concerns, doubts, and interpretive signals emerge in influencer-led social conversations that remain less visible in marketplace reviews?*

2 Literature Review

The literature on user-generated content highlights the value of online reviews in digital marketplaces, demonstrating how they reflect more than simple satisfaction or dissatisfaction. In the beauty sector, this centrality is even more pronounced, as reviews of cosmetics and skincare products help interpret practical aspects such as skin compatibility, tolerability, side effects, suitability for individual needs, and perceived risk reduction (Oh, 2020). Kim et al. (2025) emphasise the role of readability, perceived reviewer trustworthiness, and identity cues, while Jia et al. (2023) show that reviews published on marketplaces have a more immediate impact on sales than content on third-party platforms because they are closer to the moment of purchase. Overall, online reviews guide the interpretation of products characterised by high uncertainty and strong personal involvement. However, their effectiveness depends not only on the text, but on the interaction between information quality and the platform's selective mechanisms: visibility and position influence review adoption (Alzate et al., 2024), authenticity and trust signals strengthen their reliability (Liu and Dawod, 2025), while review environments impact brand loyalty and eWOM (Tran et al., 2025). Platforms like Amazon are therefore not simple repositories of opinions, but socio-technical environments that select, amplify, and make certain voices more persuasive. This is precisely where a limitation of the literature emerges: studies on marketplaces clarify helpfulness, trust, and risk reduction, but often treat the marketplace as an almost self-sufficient context, isolating it from other discursive environments that contribute to product perception (Matassi & Boczkowski, 2021; Mustak et al., 2024). From this perspective, YouTube represents a profoundly different context. Unlike marketplaces, which are dominated by review-oriented texts, YouTube creates a richer, more relational, and narrative environment. Beauty channels are perceived not only as entertainment spaces, but also as places for pre-purchase research, learning, and indirect product observation. Lee and Lee (2022) show that these channels foster both information acquisition and vicarious experiences that partially compensate for the lack of direct contact with the product. Kabel et al. (2024) suggest that spontaneous online conversations can be valuable sources for identifying needs expressed in natural language, as they offer rich, contextualised data on the current use of products and services. This literature explains influencers' persuasive mechanisms but focuses primarily on outcomes such as purchase intention, trust, and engagement. The gap in this research, therefore, lies in the fact that we know fairly well why creators are influential, but less so how evaluative and emotional signals differ between marketplace reviews and conversations generated around creator-driven content. A second line of inquiry concerns the most suitable analytical level for studying these texts. The distinction between sentiment analysis and emotion detection is crucial because it refers to two different levels of interpretation of the evaluative discourse. Traditional sentiment analysis reduces the text to a general polarity - positive, negative, or neutral - offering a synthetic measure of the overall orientation. However, this reduction is often insufficient in the beauty sector, where texts rarely express univocal judgments and combine appreciation, doubt, disappointment, curiosity, or relief with respect to different dimensions of the consumer experience. For this reason, recent literature is moving towards more granular models, capable of distinguishing not only the general tone, but also specific aspects and differentiated emotions (Aka Uymaz & Kumova Metin, 2022; Mustak et al., 2024; Mouyassir et al., 2024; Ng et al., 2024). Mouyassir et al. (2024) and Ng et al. (2024) show that transformer models improve the ability to capture this internal plurality. Emotion detection thus adds a further layer, as texts with the same valence can convey different affective states and different interpretive implications. In this context, the GoEmotions benchmark by Demszky et al. (2020), with 27 emotional categories and one neutral one, represents an important step towards finer taxonomies. Finally, the literature on cross-platform comparison shows that cross-platform comparison cannot be treated as a simple descriptive extension of analysis. Digital platforms offer different affordances, usage practices, and relational logics, which influence the production, circulation, and interpretation of content (Matassi & Boczkowski, 2021). This gives rise to the concept of analytical comparability: reviews and comments are not comparable because they are structurally identical, but because they represent different and complementary expressions of evaluation, interpretation, and reaction within the same consumption domain. The value of comparison lies in recognizing these differences as part of the object of study, not as methodological noise. The interpretative robustness of research therefore

depends not on the total elimination of biases, but on their explicitness and treatment as an integral part of the comparison itself.

3 Methodology

This study adopts a cross-platform comparative design to examine how consumer discourse differs between marketplace review environments and influencer-driven social conversations. Amazon reviews and YouTube comments serve as empirical contexts to compare two distinct forms of user-generated content. The empirical context is the beauty industry, where product evaluation is highly experiential. The analysis focuses on three active skincare products from the same brand, a choice that reduces cross-brand variability and keeps the comparison focused on platform effects and discursive environments, rather than differences in brand positioning. Amazon reviews were collected using the Apify scraper. Because the scraper imposed a limit on the number of reviews that could be retrieved, a proportional sampling strategy was adopted based on the distribution of star ratings observed on Amazon UK for each product. Amazon UK was selected to maintain linguistic consistency with the English-speaking scope of the study. For each product, a final set of 100 Amazon reviews was created, and review titles and bodies were merged into a single text field to preserve the evaluative content of each review as a single unit of analysis. The YouTube dataset was constructed through a multi-step workflow implemented in KNIME and integrated with Python. Three query formulations were used for each product: "review," "first impressions," and "wear test," as they are common in beauty discourse on YouTube and capable of capturing different product-centered video contexts. Candidate videos were retrieved via the YouTube Data API and then cleaned by removing duplicates, incomplete records, and entries with missing identifiers. Video-level metadata was added via a second API request, and only videos that met minimum quality criteria, including a minimum threshold of 50 comments, were retained. Comments were then extracted using a Python script that queried the commentThreads endpoint, managed pagination, and saved the comment text along with contextual metadata, including comment ID, author, publication date, like count, video title, video ID, and query type. Duplicate comments and records with missing text were removed before export. The study focuses on English-language content. Amazon reviews were collected from Amazon UK, while YouTube comments were filtered for English-language content, despite potentially coming from a broader set of English-speaking users. This difference is recognized as a limitation of the cross-platform design, as platform composition and audiences may partially overlap with geographic and cultural variations. Because YouTube comment sections contain many message types that are not directly comparable with marketplace reviews, all extracted YouTube comments were subjected to a preliminary ChatGPT-based classification phase. The same prompt structure was applied to all three products, modifying only the name of the focal product. Each comment was assigned to one of five unique categories: comment_creator, target_review, other_product, advice_oriented, and uncertain. The goal of this phase was to isolate comments directly relevant to the focal product and functionally comparable to Amazon reviews. Comments classified as target_review were used in the main cross-platform comparison of emotions, as they report evaluations, experiences, results, effects, or reactions to the product. Prior to validation, a gold standard was developed, defining the five categories through operational definitions, positive examples, borderline cases, and tiebreak rules. Comments reporting a clear personal experience with the focal product were assigned to target_review, while comments providing advice or asking about usage instructions, product compatibility, frequency of use, or skin type suitability were assigned to advice_oriented. The uncertain class was narrowly defined as a residual category for comments that were too short, irrelevant, incomprehensible, or impossible to reliably assign to one of the other classes. To validate the ChatGPT-based classification, a stratified sample of 300 comments was re-annotated independently by the first author using the default gold standard, without access to the model's predictions. The sample was balanced by product and by ChatGPT-predicted class, with 100 comments per product and 20 comments for each predicted class within each product.

Table 1. Validation metrics by product

product_name	n_comments	accuracy	cohen_kappa
Peeling Solution	100	0,71	0,6375
Hyaluronic Acid Serum	100	0,76	0,7
Salicylic Acid Solution	100	0,92	0,9

Table 2. Human–ChatGPT classification confusion matrix (%)

	chatgpt_comment_creator	chatgpt_target_review	chatgpt_other_product	chatgpt_advice_oriented	chatgpt_uncertain
human_comment_creator	81,25%	9,38%	3,12%	1,56%	4,69%
human_target_review	4,26%	87,23%	4,26%	2,13%	2,13%
human_other_product	0%	8,33%	76,67%	10%	5%
human_advice_oriented	5%	6,67%	10%	78,33%	0%
human_uncertain	4,35%	5,8%	5,8%	7,25%	76,81%

The validation of the ChatGPT-based classification showed an overall accuracy of 79.67% and a Cohen's κ of 0.746. Following the pre-set decision rule (κ ≥ 0.61 = substantial, Landis & Koch, 1977), the LLM-assigned labels were retained for the full corpus. The confusion matrix shows strong overall agreement between human and ChatGPT annotations, with the highest consistency for target_review, followed by comment_creator, advice_oriented, uncertain, and other_product. However, the matrix also reveals systematic ambiguities, especially between other_product and advice_oriented, where product comparisons are often interpreted as requests for guidance. Additional confusion appears between comment_creator and target_review, suggesting that comments directed at the creator may sometimes be read as evaluations of the target. The uncertain category present some misclassifications spread across advice_oriented, target_review, and other_product, indicating that vague or short comments are often forced into more specific categories by the model. The final product-level datasets underwent light text cleaning in Python. User mentions, hashtags, URLs, and redundant whitespace were removed, but no aggressive normalisation was applied. This choice was made to reduce technical noise while preserving the informal and context-dependent texture of user-generated discourse.

Emotion detection was carried out using SamLowe/roberta-base-go_emotions, a RoBERTa-based model fine-tuned on the GoEmotions benchmark. Instead of retaining only the first predicted emotion, the model was run with top_k=None in order to extract the full score distribution across the 28 GoEmotions labels. For each comment, the pipeline stored the scores for all 28 GoEmotions labels, the top three emotions, and their associated model scores. The full emotion-score distribution was then aggregated into five macro-emotional categories: positive, negative, uncertainty, cognitive, and neutral. The positive category includes emotions such as admiration, approval, gratitude, joy, love, optimism, and relief. The negative category includes anger, annoyance, disappointment, disapproval, disgust, fear, nervousness, remorse, and sadness. The uncertainty category includes confusion and curiosity. The cognitive category includes realization and surprise. The neutral category includes the neutral label. For each comment, macro-emotion scores were computed by summing the scores of the emotions belonging to each macro-category. The dominant macro-emotion was also identified as the macro-category with the highest aggregated score. However, the main comparative analysis relies on the weighted macro-emotion distribution rather than only on the dominant label, since weighted scores preserve the multidimensional emotional profile of each comment. Because GoEmotions was fine-tuned on Reddit comments, model performance on Amazon and YouTube text is not guaranteed. To validate the classifier in-domain, we manually re-annotated a stratified random sample of 150 valid texts. We report Cohen's κ between the model and the human gold standard at the five macro-category level (Positive / Negative / Cognitive / Uncertainty / Neutral), κ = 0.75 (accuracy = 0.83). Since the validation supports more reliable agreement at the macro-category level, the inferential analysis in §4 is conducted at the macro-category level, while fine-grained results are presented as descriptive.

4 Results

4.1 Differences in weighted macro-emotions across platforms

The aggregate distribution of weighted macro-emotions shows a clear difference between marketplace reviews and influencer-led conversations. Amazon reviews are strongly concentrated around positive emotions, which represent approximately two-thirds of the overall weighted emotional profile. YouTube comments, on the other hand, exhibit a more distributed emotional structure: positive emotions remain present, but have a much lower relative weight, while negative, uncertainty, and neutral components increase.

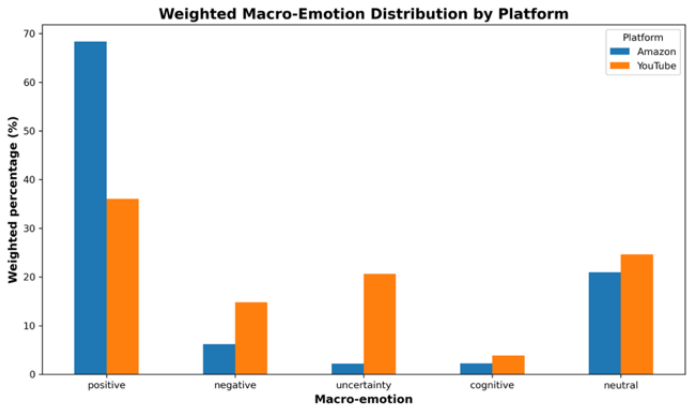


Figure 1 – Weighted Macro-Emotion Distribution by Platform

4.2 Product-level robustness check

The product-level heatmap confirms that a single product does not drive the aggregate pattern. In all three cases analysed, YouTube comments show a lower positive component than Amazon reviews. The difference is particularly marked for Peeling Solution, followed by Hyaluronic Acid and Salicylic Acid. At the same time, YouTube shows higher values for the negative components for all products and higher values for uncertainty, especially for Peeling Solution, potentially associated with concerns about frequency of application, skin sensitivity, irritation, and compatibility with other products. These results suggest that influencer-led conversations become particularly informative when the product requires interpretation, caution, or practical guidance. The product-by-product comparison reinforces the aggregate result, but also shows that the intensity of the difference between platforms varies based on product characteristics and perceived risk level.

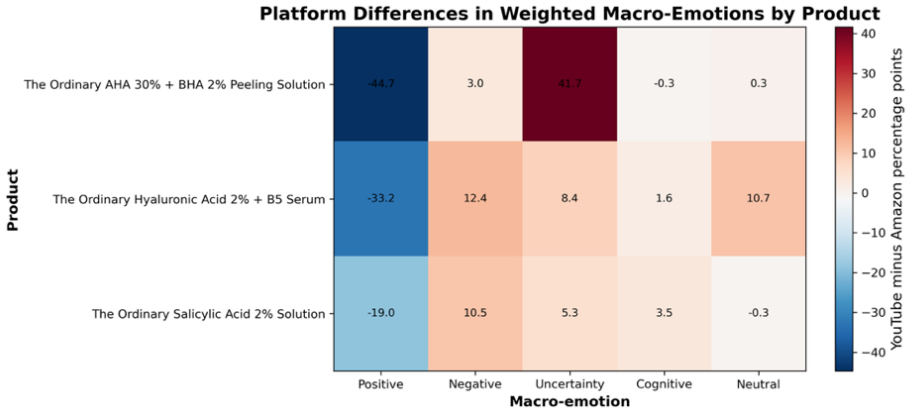


Figure 2 – Platform Differences in weighted Macro-Emotions by Product

4.3 Specific emotions and emotional ambiguity

The distribution of the top ten emotions helps better interpret the results observed at the macro-category level. Amazon reviews are strongly driven by positive emotions, particularly admiration and, to a lesser extent, approval. This confirms that marketplace reviews tend to express direct appreciation and more concentrated evaluative judgments. YouTube comments, instead, are less concentrated around a single positive emotion. The neutral category is particularly visible, as is curiosity, along with other emotions such as confusion, sadness, gratitude, disapproval, and love.

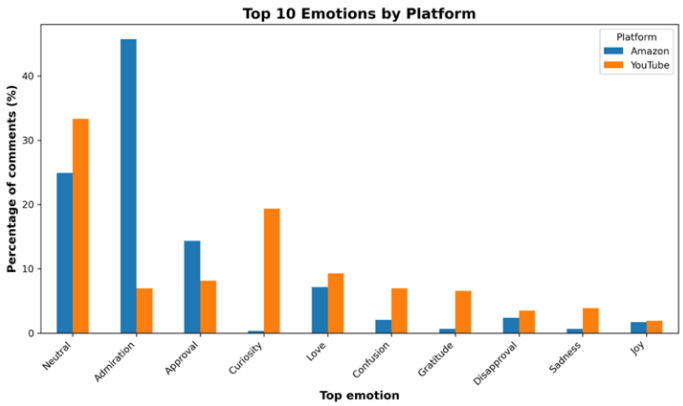


Figure 3 – Top 10 Emotions by Platform

Further evidence emerges from the average margin between the first and second emotions detected by the model. Amazon reviews show a higher average margin, indicating that RoBERTa more clearly identifies a dominant emotion. YouTube comments, however, show a lower average margin, suggesting greater emotional ambiguity or the coexistence of different emotional signals.

5 Discussion and Conclusions

This study compared marketplace reviews and influencer-led social conversations to understand whether these two digital environments reveal different emotional profiles and information needs in the context of beauty products. The results confirm that the two sources are not interchangeable: even when the comparison is based on balanced samples and YouTube comments filtered for product relevance, marketplace reviews and influencer-led conversations display different emotional structures. In answer to RQ1, regarding the differences between emotional profiles in the two environments, the results show that marketplace reviews are characterized by a greater concentration of positive emotions. Emotional expression on Amazon appears more stable, evaluative, and oriented toward synthetic product judgments. In contrast, YouTube comments display a more heterogeneous emotional profile, with a lower concentration of positive emotions and a greater presence of negative components, uncertainty, and cognitive processing. In response to RQ2, regarding additional signals emerging in influencer-led conversations, the results indicate that YouTube makes doubts, requests for clarification, uncertainties about use, and weak signals related to practical experience with the product more visible. This element emerges both from the emotional analysis, in which uncertainty and cognitive components are more evident in YouTube comments, and from the preliminary classification of comments, which identified a significant presence of content classifiable as advice-seeking. From a theoretical perspective, the study contributes to the literature on user-generated content, online reviews, and influencer ecosystems by showing that platforms are not simply neutral containers of opinions, but discursive environments that shape consumer experience. Marketplace reviews tend to make the final judgment observable, often expressed in a more concise and evaluative form. Influencer-led conversations, on the other hand, allow us to observe the process through which users interpret, negotiate, and contextualize product use. The methodological contribution involves the construction of a comparative and validated pipeline for analyzing heterogeneous content from different platforms. The preliminary classification of YouTube comments using ChatGPT was validated against a human gold standard, showing substantial agreement with human annotation. The emotional classification using RoBERTa-GoEmotions was also validated at the macro-category level, explicitly acknowledging the limitations of transferring a model trained on Reddit to different domains such as Amazon and YouTube. From a managerial perspective, the findings suggest that beauty companies should not rely solely on marketplace reviews to understand consumer perceptions. The latter are useful for monitoring satisfaction, approval, and summary product reviews, but they may be less effective in identifying emerging concerns, difficulties with use, or unclear information needs. Influencer-led social conversations, however, can serve as a particularly useful social listening space for identifying recurring questions, concerns and requests for product education. These insights can support brand communication, the development of FAQs, the selection of content to be shared with creators, and the improvement of instructions for use.

The study has some limitations that should guide the interpretation of the results. The analysis used only one brand and three products. The final comparison is based on balanced samples, not the whole corpus. Only English YouTube comments and Amazon UK reviews directly related to the product were included. The emotion

detection model was applied outside its training context. Thus, these results show cross-platform differences within a specific setting, not universal emotional signatures. Future research could include a program trained for the specific communication environment, more brands, larger samples, other product categories, or the analysis of excluded categories, such as advice-seeking comments.

Despite these limitations, the study shows that comparing platforms allows for a richer understanding of consumers' digital discourse and, in particular, that influencer-led environment make visible dimensions of product interpretation that are only partially observable in brand-led marketplace environment.

6 Ethics Declaration

Ethical clearance was not required for this research, as the study relied exclusively on publicly available online user-generated content and did not involve direct interaction with human participants. The data were analysed in aggregated form, and no personal identifying information is reported in the paper.

7 AI Declaration

AI tools were used to support the preliminary classification of YouTube comments into analytical categories, as described in the methodology section. AI tools were also used to support language editing and reference-format checking. All outputs were reviewed by the authors, who remain fully responsible for the research design, analysis, interpretation of results, and final content of the paper.

References

Uymaz, H. and Kumova Metin, S. (2022) 'Vector based sentiment and emotion analysis from text: a survey', *Engineering Applications of Artificial Intelligence*, 113, 104922. doi: 10.1016/j.engappai.2022.104922.

Alzate, M., Arce-Urriza, M. and Cebollada, J. (2024) 'Is review visibility fostering helpful votes? The role of review rank and review characteristics in the adoption of information', *Computers in Human Behavior*, 153, 108088. doi: 10.1016/j.chb.2023.108088.

Demszky, D., Movshovitz-Attias, D., Ko, J., Cowen, A., Nemade, G. and Ravi, S. (2020) 'GoEmotions: a dataset of fine-grained emotions', in *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics, pp. 4040–4054. doi: 10.18653/v1/2020.acl-main.372.

Jia, Y., Feng, H., Wang, X. and Alvarado, M. (2023) '"Customer reviews or vlogger reviews?" The impact of cross-platform UGC on the sales of experiential products on e-commerce platforms', *Journal of Theoretical and Applied Electronic Commerce Research*, 18(3), pp. 1257–1282. doi: 10.3390/jtaer18030064

Kabel, D., Martin, J., Elg, M. and Witell, L. (2024) 'Capturing the voice of the customer: focus groups versus netnography?', *Total Quality Management & Business Excellence*, 35(11–12), pp. 1359–1377. doi: 10.1080/14783363.2024.2369592.

Kim, S., Park, S., Li, X. and Kim, J.K. (2025) 'A deep learning-based review helpfulness prediction system for online beauty products', *IEEE Access*, 13, pp. 202049–202061. doi: 10.1109/ACCESS.2025.3638451

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174. Doi: https://doi.org/10.2307/2529310

Lee, M. and Lee, H.-H. (2022) 'Do parasocial interactions and vicarious experiences in the beauty YouTube channels promote consumer purchase intention?', *International Journal of Consumer Studies*, 46(1), pp. 235–248. doi: 10.1111/ijcs.12667.

Liu, X. and Yahya Dawod, A. (2025) 'When technology signals trust: blockchain vs. traditional cues in cross-border cosmetic e-commerce', *Information*, 16(10), 913. doi: 10.3390/info16100913.

Mahbub, F.B., Mim, M.H. and Rahman, M.S. (2024) 'Sentiment analysis of beauty product reviews in Bangladesh: a comprehensive study using transformer-based models', in *2024 27th International Conference on Computer and Information Technology (ICCIT)*. IEEE, pp. 2712–2717. doi: 10.1109/ICCIT64611.2024.11021928.

Matassi, M., & Boczkowski, P.J. (2021). An Agenda for Comparative Social Media Studies: The Value of Understanding Practices From Cross-National, Cross-Media, and Cross-Platform Perspectives.

Mouyassir, K., Fathi, A. and Assad, N. (2024) 'Elevating aspect-based sentiment analysis in the Moroccan cosmetics industry with transformer-based models', *International Journal of Advanced Computer Science and Applications*, 15(6). doi: 10.14569/IJACSA.2024.0150654.

Mustak, M., Hallikainen, H., Laukkanen, T., Plé, L., Hollebeek, L.D. and Aleem, M. (2024) 'Using machine learning to develop customer insights from user-generated content', *Journal of Retailing and Consumer Services*, 81, 104034. doi: 10.1016/j.jretconser.2024.104034.

Ng, C.M., Tiun, S. and Sastria, G. (2024) 'Multi-label aspect-sentiment classification on Indonesian cosmetic product reviews with IndoBERT model', *International Journal of Advanced Computer Science and Applications*, 15(11). doi: 10.14569/IJACSA.2024.0151168.

Oh, Y.-K. (2020) 'Determinants of online review helpfulness for Korean skincare products in online retailing', *Journal of Distribution Science*, 18(10), pp. 65–75. doi: 10.15722/jds.18.10.202010.65.

Tran, S.N., Trương Bich, P., Doan, T.L.M., Lam, V.L.C. and Tran Thi Anh, T. (2025) 'The impact of online reviews on brand loyalty and e-WOM in Vietnam's cosmetics e-commerce sector: the moderating role of e-service quality', *Social Sciences & Humanities Open*, 12, 102208. doi: 10.1016/j.ssaho.2025.102208.

Metrics and Value of Intelligent Knowledge: How Organizations Turn AI Options into Decision Capability, and Become More Sustainable

Harald M. Grundner
innoVAVE, Hockenheim, Germany

Abstract

AI, copilots and autonomous agents have drastically lowered the cost of producing information and alternatives. Yet many organizations do not become faster or more effective in decision-making. Instead, they experience option overload, longer alignment cycles and persistent uncertainty at the decisive moment. This paper argues that the bottleneck has shifted from generating knowledge artefacts to building organizational decision capability.

The practitioner message is simple: value is always two-sided: *customer benefit and economic gain for the organization*. Intelligent knowledge is therefore only valuable when it enables better decisions that increase customer benefit and improve value capture. Sustainability in day-to-day management becomes tangible when results and artefacts are reusable: a product (material or immaterial, such as a copilot, agent or prompt) becomes sustainable only if its use repeatedly creates value for users and customers.

We define intelligent knowledge pragmatically as knowledge that makes decisions actionable, accountable and auditable. To measure its value, we propose a metric architecture with three layers. Outcome metrics ask whether decisions improved results and reduced rework and waste. Flow metrics ask whether the right knowledge reached decisions in time and whether it became reusable as a standard or building block. Leadership and governance metrics ask whether decisions are owned, traceable and trustworthy, including disciplined use of AI.

To operationalize the framework, we embed it in a recurring rhythm: Learn – Measure – Build/Test – Decide. Each leaf ends with a decision gate. New insights can force a sharp return and re-entry, ensuring learning is woven back into the process. Each cycle yields two outputs: a decision-ready set for the current decision episode, and a process result that updates knowledge assets (standards, templates, prompts, playbooks) for continuous improvement. An illustrative scenario shows how organizations can transform AI-driven option abundance into calmer, faster and more sustainable decisions by making value criteria explicit and by measuring both decision impact and reusability.

Keywords: Intelligent knowledge; decision capability; knowledge metrics; AI governance; reusability

Paper type: Practical Paper

1 AI Scales Options – Organizations Need Decision Capability

We are witnessing a shift that many still underestimate. In the past, knowledge was expensive, slow, and often difficult to access. Today AI can compress information within seconds, generate alternatives, structure arguments and produce recommendations. The cost of producing knowledge artefacts has collapsed. Yet decisions often do not become easier; they become harder. This is rarely a technology issue. It is a leadership and organizational issue.

The practitioner's anchor is value. Value is not a slogan. Value is a test: customer benefit and economic gain for the organization. If knowledge, a copilot, an agent or a prompt does not improve customer benefit and value capture, it is activity not value. And if the result is not reusable, it is not sustainable in practice. Sustainability becomes tangible when value becomes repeatable: reuse is the day-to-day signal that an artefact creates value more than once.

Decision theory reminds us why this matters. Organizations decide under time pressure, conflicting goals and bounded rationality (Simon, 1947/1997; March, 1994). AI does not remove this reality; it intensifies it by increasing the number of possible paths. AI scales speed and options, but it does not scale responsibility or the capability to learn, test and decide consciously.

In practice, this mismatch creates a familiar pattern: more dashboards, more slideware, more 'evidence', and still a foggy decision meeting. The missing piece is not information. The missing piece is decision capability. The ability to translate knowledge into a decision that is owned, explainable, and resilient to volatility.

This paper provides a measurement and leadership logic to translate AI outputs into decision capability through three metric layers—outcome, flow, and leadership/governance—and an operating rhythm of Learn – Measure – Build/Test – Decide. The aim is not to add reporting. The aim is to make decision quality measurable and improvable, so that knowledge and AI support repeatedly create customer benefit and economic gain.

2 Knowledge Becomes Intelligent In Decision-Making

This paper is not about building larger repositories or optimizing AI performance in isolation. The core is decision capability. The organizational ability to create orientation under uncertainty, make assumptions explicit, test what matters and decide with ownership. The value of knowledge assets and AI artefacts is assessed by their contribution to decision episodes: real choices with consequences.

To avoid confusion, we distinguish three objects. Knowledge assets include documents, standards, lessons learned, routines and tacit expertise. AI artefacts include copilots, agents, models and analytics pipelines. Decision episodes are the moments where options are closed and commitments are made. Knowledge and AI matter because they influence those decision episodes, either by reducing uncertainty or by creating false certainty.

Sustainability becomes concrete when it is linked to reuse. A knowledge asset is sustainable when it can be reused and extended as a standard or building block, because it repeatedly creates value. An AI artefact becomes sustainable when it is used repeatedly in a way that produces valuable outcomes for users and customers, without eroding trust.

This perspective reframes “intelligence”. Intelligence is not a property of a tool alone. It is the quality of the socio-technical system: how people, knowledge and AI work together to produce decisions. In that sense, intelligent knowledge is an organizational capability, not an IT feature.

3 Methodological Positioning: A Practice Framework With Design Intent

This contribution is a practice-oriented framework proposal with design intent. It synthesizes practitioner patterns with selective scientific grounding: bounded rationality and decision processes (Simon; March), knowledge as strategic resource and social process (Grant, 1996; Nonaka and Takeuchi, 1995; Tsoukas, 1996), and governance frameworks for trustworthy AI (NIST, 2023; OECD, 2019/updated; ISO 30401; EU AI Act context).

Design intent means the framework is formulated as an artefact that can be applied, adapted and evaluated. It specifies what should be measured (value realized, value repeatability, and value responsibility), where it should be measured (decision episodes), and how it becomes operational (Cloverleaf rhythm and SIPOC+ workflow).

The framework is conceptual, but not detached. It is grounded in recurring field patterns: decision delays despite abundant information, late rework caused by premature lock-in, unclear ownership that inflates coordination cost, and AI-driven option overload that creates the illusion of progress while decisions remain fragile.

A brief application scenario is used as an illustrative operationalization artefact rather than an empirical proof. At the same time, the framework is designed to be evaluable in future work. A plausible evaluation approach would trace decision episodes before and after adoption: decision cycle time, late-change load, trace completeness, reuse rates of standards or prompt modules, and perceptions of decision clarity and trust.

4 Intelligent Knowledge: Actionable, Accountable, Auditable – And Value-Positive

We define intelligent knowledge pragmatically: knowledge is intelligent when it makes decisions actionable, accountable and auditable. And Knowledge is intelligent when it is value-positive in the practitioner sense: it increases customer benefit and improves value capture.

Actionable means we move from ‘interesting’ to ‘we act’. It is not the amount of information that creates actionability; it is the quality of selection and constraint. In AI-rich environments, actionability often improves when teams explicitly define what they will not pursue and why, because option overload is the enemy of action.

Accountable means responsibility does not disappear into team fog. It means a decision owner is visible, and the organization knows who carries consequences. Accountability is not bureaucracy; it is speed. Where ownership is unclear, decisions turn into endless alignment. Where ownership is clear, debate becomes preparation and decision becomes execution.

Auditable means assumptions, evidence and rationale can later be reconstructed. This supports learning, but it also supports trustworthy AI use. If a decision rationale cannot be reconstructed, the organization loses two things at once: trust and learning. It repeats discussions, redoes analyses, and loses tempo.

The issue is not that we ‘build wrong’. The issue is premature lock-in. That means turning assumptions too early into architecture, specification and commitments, making correction expensive and resource-intensive.

That is where sustainability becomes tangible: fewer premature lock-ins, earlier correction, and less waste of time, resources and trust.

5 Three Metric Layers That Pull The Cycle Toward Value And Reuse

The metric architecture distinguishes three layers. Each layer has two roles: planning (targets) and evaluation (criteria). This dual role matters because metrics should not only judge results after the fact; they should steer the cycle while it runs.

Outcome metrics plan and assess value realized. They make explicit what "better" means for the decision episode: improved customer benefit, improved value capture, fewer late changes, and less rework and waste. *Flow metrics* plan and assess value repeatability. They ensure the right knowledge reaches the decision in time and that the cycle produces reusable standards, prompts, templates or building blocks. *Leadership and governance metrics* plan and assess value responsibility. They ensure decisions are owned and traceable, and AI is used in a disciplined and trustworthy way.

Together, the three layers prevent a common failure mode: organizations measure activity (documents, meetings, prompt usage) while decision quality silently deteriorates. A balanced combination is required: outcome without flow becomes coincidence; flow without governance becomes risk; governance without outcome becomes bureaucracy.

Outcome metrics answer a direct value question: did this cycle improve results and reduce rework and waste? In planning, outcome targets can include higher decision stability, lower late-change load, and measurable improvements in customer benefit or economic contribution. In evaluation, outcome metrics ask whether these effects happened early enough to matter.

Practically, three outcome indicators are often powerful and easy to observe. First, late-change load: how many significant changes occur after a point of commitment. Second, decision stability: how often a decision is revisited because evidence was weak, assumptions were hidden, or ownership was unclear. Third, rework intensity: effort spent correcting avoidable mistakes caused by premature lock-in. These indicators do not replace financial measures; they precede them. Financial results are lagging. Decision quality is the early signal.

A typical outcome blind spot is to equate value with immediate financial metrics only. In early phases, value often appears as learning that prevents expensive detours. If knowledge helps a team avoid a wrong architecture choice, the value is real even before revenue is realized. Outcome metrics should therefore include leading indicators that connect directly to decision quality.

Flow metrics make sustainability measurable. If sustainability means repeatable value, then reuse is the practical indicator that value is not a one-off accident. In planning, flow targets can include shorter time-to-knowledge, higher reuse rate of validated rationales, and better freshness discipline. In evaluation, flow metrics ask: did the right knowledge reach the decision, and did we capture what we learned so it can be reused?

A simple flow indicator is time-to-knowledge: how long it takes to locate decision-relevant information with sufficient trust. Another is reuse-rate: how often standards, templates, prompt modules, or proven decision rationales are reused without rework. A third is assumption transparency: how often assumptions are explicit and testable rather than implicit and politically defended.

Flow metrics also protect against a typical AI failure mode: the production of plausible text that is detached from context, versioning and evidence. When AI outputs are treated as knowledge without provenance, flow looks high, but trust erodes. Flow metrics should therefore include traceability elements such as version clarity, source anchoring, and feedback closure. Did we record what happened after the decision and fold it back into the knowledge base?

Leadership and governance metrics ensure that value is created responsibly and fast. In planning, governance targets include visible decision ownership, trace completeness, and disciplined use of AI support. In evaluation, governance metrics ask whether ownership was real (not symbolic), whether rationale is reconstructable, and whether AI outputs were challenged appropriately.

A key governance indicator is trace completeness: can a third party reconstruct what was assumed, tested and decided? Another is override discipline: when humans override AI recommendations, is the rationale documented and learned from, or does override remain a private act? This is not bureaucracy. It is the mechanism that prevents 'confidence theatre' and enables systematic learning.

In practice, governance metrics also reduce coordination cost. When rationale and ownership are clear, teams spend less time re-arguing old issues. Governance becomes the steering wheel: it keeps the cycle moving clockwise with speed and trust.

Governance metrics are often misunderstood as compliance. The practitioner view is different: governance is a speed enabler because it prevents re-litigation. A decision that is accountable and auditable is less likely to

be reopened later. It reduces the need for defensive documentation. It also supports trustworthy AI use: AI outputs are accepted as hypotheses that must pass explicit gates, rather than as authoritative answers.

For flow metrics, it is useful to distinguish between 'reuse as convenience' and 'reuse as strategy'. Convenience reuse means copying a document. Strategic reuse means reusing a validated rationale, a tested option comparison method, a stable prompt module, or a building block in a modular architecture. Strategic reuse has higher leverage: it increases speed and decision quality simultaneously. Flow metrics should therefore reward strategic reuse, not only copying.

Another practical technique is to separate value creation from value capture explicitly. Customer benefit is not automatically captured as profit; profit is not automatically sustainable. Outcome metrics should therefore include at least one customer-facing indicator (benefit in use, friction reduced, adoption readiness) and one organization-facing indicator (effort avoided, rework reduced, time saved, or risk reduced). This keeps the two-sided value definition alive throughout the cycle.

A practical technique is to treat targets as a contract for speed. Teams often believe that more evidence will always produce better decisions. In reality, the best decisions are those made with sufficient evidence at the right time. Targets define what 'sufficient' means: how much uncertainty is acceptable at this stage, what must be tested now, what can be postponed, and what must never be postponed. This is where the planning role of metrics becomes leadership.

Targets should be formulated as directional improvements rather than as pseudo-precision. For example, "reduce late-change load" is a better starting point than forcing a single number that will be negotiated away. Over time, targets can mature into explicit thresholds, e.g. an agreed trace completeness threshold for decision readiness, or a reuse rate target for templates and prompt modules. The key is consistency: teams should not change the game mid-cycle.

Metric selection should be conservative. In practice, the goal is not to cover everything. The goal is to make the cycle steerable. A small number of well-chosen measures can change behavior more effectively than a large catalogue. A useful rule is to select metrics that (a) are visible in the decision episode, (b) can be influenced by the team during the cycle, and (c) have a clear value interpretation in terms of customer benefit or economic gain.

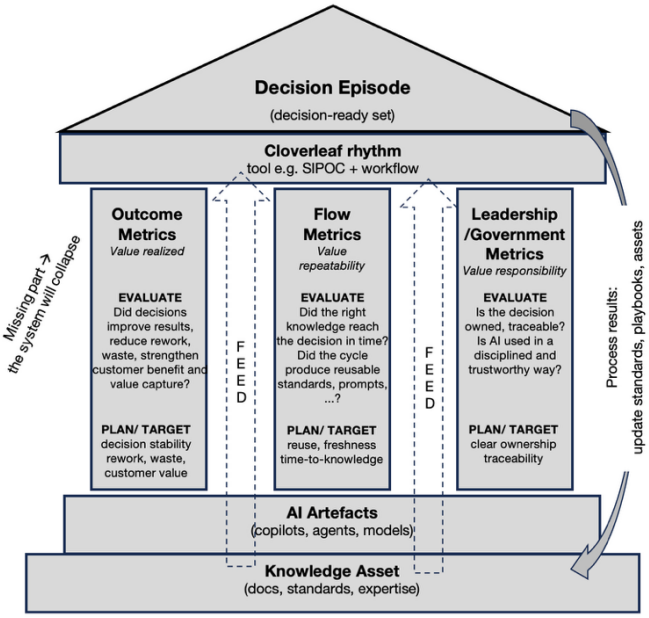


Figure 1. Metrics house: foundation (knowledge assets + AI artefacts), three metric pillars (plan + evaluation), operationalization cornice, and decision episode roof with learning loop back into assets.

6 Learn – Measure – Build/Test – Decide: Cloverleaf Rhythm With Decision Gates

To become everyday reality, the framework is embedded in a leadership rhythm: *Learn – Measure – Build/Test – Decide*. The rhythm is intentionally simple, but it is not linear. Each leaf ends with a decision gate. The gate is the discipline point where a team decides whether it can continue clockwise, or whether new insight forces a sharp return and re-entry.

This “hairpin return” is not failure. It is the mechanism that prevents premature lock-in. Without it, teams move forward with the wrong certainty: they build faster, but they build commitments on unstable assumptions. With it, learning is woven back into the process and the cycle regains traction.

Each gate should be expressed as a value question, not only a process question. Leaf 1 (Learn) asks: do we truly understand customer benefit, or are we only capturing requirements? Leaf 2 (Measure) asks: can we substantiate benefit, cost and impact, or are we merely claiming them? Leaf 3 (Build/Test) asks: did users actually experience the benefit, and what does it really cost? Leaf 4 (Decide) asks: is the decision value-positive, and are we making it reusable? These questions keep the cycle tied to metrics and value.

Importantly, each cycle must yield two outputs. Output A is the decision-ready set for the current decision episode: criteria, evidence, risks, trade-offs, and a recommended commitment with ownership. Output B is the process result: reusable standards, templates, playbooks and prompt modules that strengthen the foundation for the next cycle. With Output B, sustainability becomes tangible: value becomes repeatable through reuse. Without Output B, organizations may deliver once, but they do not improve.

This dual-output logic is also what turns AI tools into sustainable assets. A copilot or agent becomes sustainable when its usage pattern stabilizes into reusable modules: prompt patterns, evaluation checklists, trace templates, and decision rationales. The organization shifts from asking AI for answers to building decision-ready knowledge assets with AI support.

In AI-enabled work, Output B is also where tools become assets. A prompt that helped once becomes a reusable prompt module when it includes: purpose, input assumptions, evaluation criteria, and a trace note on when it should not be used. An agent becomes a reusable building block when it has guardrails: clear boundaries, escalation rules, and a traceable rationale for its recommendations. This turns AI use from ad-hoc interaction into repeatable value creation.

The dual-output logic is the bridge between project work and organizational learning. Output A serves the current decision episode. Output B serves future decision episodes. In practice, Output B should be assigned an owner and a short completion rule: the cycle is not closed until the reusable artefact is stored in a findable place, linked to its context, and tagged as reusable. Otherwise, learning remains local.

The hairpin return becomes valuable when it is socially safe. Teams must be allowed to go back without being accused of ‘losing time’. In practice, this requires leadership language: returning is not failure; returning is avoiding waste. The faster a team can admit “we learned something that changes our direction,” the more sustainable it becomes, because it avoids expensive late corrections.

To make the gates operational, teams should define what is expected at each gate in one sentence. The sentence is not a checklist; it is a boundary. For instance: ‘We can continue when customer benefit is expressed as an outcome in use.’ Or: ‘We can continue when at least one critical assumption has been tested with evidence that would change our mind.’ Such boundary sentences make the gate real.

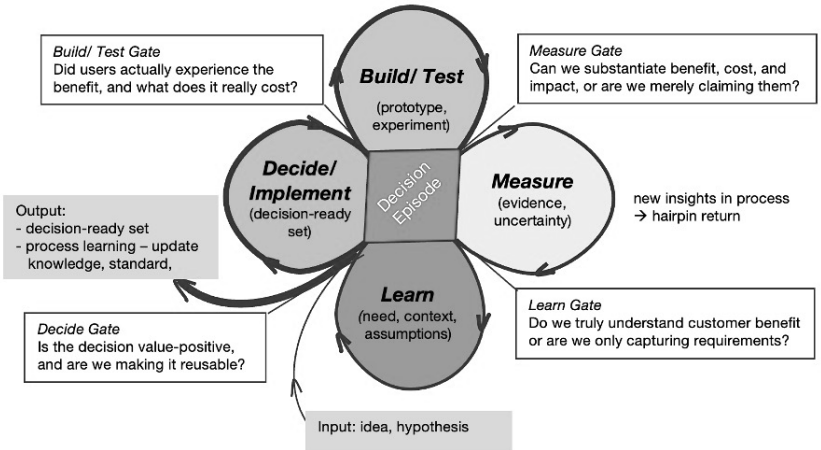


Figure 2. Cloverleaf logic (clockwise): decision gates, sharp return on new insight, and dual outputs (decision-ready set + continuous improvement into knowledge assets).

7 Application Scenario: Turning Option Abundance Into Calm Decisions

A typical example is a development project in a volatile market. Customers themselves do not yet know precisely what they will truly need in twelve months. Cost and sustainability requirements are high. Supply

chains are fragile. At the same time, the organization introduces a copilot that produces alternatives and justifications within seconds. Material explodes, yet decisions slow down.

The reason is missing decision capability. Teams accumulate options but do not close them. Assumptions remain implicit. Criteria are unclear or disputed. Responsibility is diffused. AI then amplifies the pattern: each function can generate its own 'evidence', and meetings become a battle of outputs rather than a synthesis of insight.

Applying the approach begins with defining critical decision episodes. For example, an architecture choice is framed as a decision episode with explicit criteria: customer benefit, total cost, risk exposure, and sustainability impact. The team then structures decision-relevant knowledge: the key assumptions behind each option, what evidence exists, what must be measured, what must be tested, and who owns the final commitment.

Next, the Cloverleaf rhythm is used to drive progress. In Learn, the team clarifies customer benefit beyond requirements by identifying what users are trying to achieve and what trade-offs they accept. In Measure, the team turns claims into testable hypotheses and selects a small set of measures that reduce uncertainty fastest. In Build/Test, prototypes and experiments are run, not to prove everything, but to invalidate weak assumptions early. In Decide, the team commits with ownership and a traceable rationale.

The metric layers act as guardrails. Outcome targets for the architecture decision include reduced late-change load and improved stability of key choices. Flow targets include time-to-knowledge for similar past decisions and reuse of a validated architecture rationale template. Governance targets include clear ownership and trace completeness. During the cycle, the evaluation questions are used in the gates: did we improve results and reduce rework/waste; did the right knowledge reach the decision and become reusable; is the decision accountable and auditable?

The practical effects are usually visible quickly. Discussions become shorter because assumptions are explicit. Conflicts become more factual because criteria are agreed. Decisions become calmer because they are justified by evidence rather than defended by authority or gut feeling. Sustainability becomes tangible because the cycle produces reusable assets: a standard decision template, a prompt module for option comparison, a trace checklist, and a small playbook for future architecture decisions. That process output reduces effort in the next project and increases the probability of value creation repeating.

A third shift is the disciplined capture of reuse. The team produces a decision-ready set and, in parallel, packages the process result into assets: a decision rationale template, a prompt module library entry, a trace checklist, and a short playbook. This is where sustainability becomes measurable. In the next project, time-to-knowledge drops because the team does not start from scratch. The organization becomes faster and more consistent, not because it works harder, but because it reuses what it learned.

A second shift is the explicit handling of uncertainty. The team agrees on a limited number of uncertainty drivers and treats them as the 'learning agenda' of the cycle. Measures and tests are chosen to reduce those uncertainties quickly. This avoids the typical trap of measuring what is easy. It also keeps the process fast because measurement is not a reporting activity; it is a decision activity.

In the scenario, the first visible shift is the quality of questions. Instead of asking the copilot for "the best option," the team asks for option sets framed by value. The prompt is shaped by the gates: customer benefit definition, measurable claims, test design, and decision readiness. This changes the entire dynamic: the copilot becomes an accelerator for structured thinking rather than a generator of plausible narratives.

Supplier	Input	Process step (activity)	Output	Customer	Targets			Cloverleaf stage and questions
					Outcome	Flow	Governance	
Project Manager	Phase result (ideas, risks, status)	Synthesize + prepare gate package (benefit definition, assumptions, open uncertainties, proposed targets)	Decision proposal (release Measure) + trace stub (assumptions& rationale draft)	Decision owner	expected higher decision stability; lower late-change exposure	time-to-knowledge target; reuse candidate identified (standard/ module)	decision owner confirmed; trace structure defined	**Learn** Do we truly understand customer benefits (not only requirements)?
Decision owner	Decision proposal (release **Measure**)	Gate decision + document rationale + release **Measure** + create learning asset seed	Measure leaf release + hypotheses charter (testable hypothesis + measures) + asset backlog entry	Project Manager/ Team	did we improve results, reduce rework-waste enough to proceed?	did the right knowledge reach the decision in time, and did we define reuse deliverables?	is the decision accountable and auditable (assumptions, evidence, plan, ownership)?	**Learn Gate → Measure Release – definition of testable hypothesizes; set of measures**

Figure 3. SIPOC+ decision template linking roles, process steps, outputs and metrics to the Cloverleaf cycle (excerpt).

8 Conclusion: Value, Reuse, And Decision Capability

AI makes many things possible, but value is created by decisions not by possibilities. Value is two-sided: customer benefit and economic gain. Sustainability becomes tangible when value is repeatable through reuse of knowledge assets, prompts, copilots, agents and standards.

Intelligent knowledge is therefore decision capability gained: actionable, accountable and auditable decisions that improve customer benefit, improve value capture, and leave behind reusable assets that strengthen the next cycle. In practice, this means each cycle must produce two outputs: the decision-ready set and the process learning captured as assets. That is how organizations become faster without blind flight, and more sustainable by reducing premature lock-in and waste.

Future work can evaluate the framework through decision-episode tracing in case settings, comparing decision cycle time, late-change load, trace completeness and reuse rates before and after adoption. The key test remains practical: does the organization repeatedly create customer benefit and economic gain, and does it turn learning into reusable assets?

9 Ethics Declaration

No ethical clearance was required for this conceptual/practice-oriented framework paper; no human subject data was collected.

10 AI Declaration

AI tools were used as writing and editing support to draft, translate and structure text. The author reviewed, curated and revised all content and takes full responsibility for the final manuscript.

References

Argyris, C. and Schön, D.A. (1978) Organizational Learning: A Theory of Action Perspective. Reading, MA: Addison-Wesley.

European Commission (2024) AI Act enters into force. Brussels.

Grant, R.M. (1996) 'Toward a knowledge-based theory of the firm', Strategic Management Journal, 17(S2), pp. 109–122.

ISO (2018) ISO 30401:2018 Knowledge management systems — Requirements. Geneva: International Organization for Standardization.

March, J.G. (1994) A Primer on Decision Making: How Decisions Happen. New York: Free Press.

NIST (2023) Artificial Intelligence Risk Management Framework (AI RMF 1.0). Gaithersburg, MD: National Institute of Standards and Technology.

Nonaka, I. and Takeuchi, H. (1995) The Knowledge-Creating Company. New York: Oxford University Press.

OECD (2019/updated) OECD AI Principles. Paris: OECD.

Simon, H.A. (1997) Administrative Behavior. 4th edn. New York: Free Press. (Original work published 1947).

Spender, J.-C. (1996) 'Making knowledge the basis of a dynamic theory of the firm', Strategic Management Journal, 17(S2), pp. 45–62.

Tsoukas, H. (1996) 'The firm as a distributed knowledge system', Strategic Management Journal, 17(S2), pp. 11–25.

Artificial Intelligence and Labour Law: The Employers' Obligation to Reskill Workers in the Digital Era

Antonio Capasso, Annabella Lattarulo
Universitas Mercatorum, Rome, Italy

Abstract

Artificial Intelligence is driving unprecedented structural reconfiguration of labour markets, yet regulatory frameworks governing employment relationships remain anchored to twentieth-century industrial paradigms. This paper addresses a critical research gap: the absence of coherent legal frameworks allocating responsibility for workers' reskilling among employers, workers, and public institutions in contexts of AI-driven technological transformation. Drawing on comparative legal analysis and organizational learning theory, this study examines how jurisdictions—specifically Italy, the European Union, the United States, and selected developing countries—approach the regulatory challenge of employer-mandated reskilling obligations. The paper proposes a multidimensional regulatory framework articulated along five core dimensions: legal recognition of reskilling as an employment relationship component, cost-sharing mechanisms, algorithmic oversight, collective bargaining strengthening, and corporate due diligence integration. Through this framework, the paper contributes theoretically to understanding the intersection of technological innovation, labour law, and organizational learning, while offering practical recommendations for legislators and international organizations navigating AI governance in employment contexts. The analysis reveals that while the European Union has advanced more structured approaches through the AI Act, significant implementation inconsistencies persist, and integrations with labour law remain underdeveloped. Conversely, the United States exhibits fragmentation across state-level frameworks, while developing countries face structural barriers to coherent reskilling policies. The paper demonstrates that sustainable AI governance requires reconceptualizing the social contract underlying employment relationships, ensuring that technological transitions develop in alignment with adequate worker protection and organizational learning capacity.

Keywords: Artificial Intelligence, Labour Law, Reskilling Obligations, Organizational Learning, Digital Transformation, Regulatory Framework, Comparative Analysis

Paper type: Academic Research Paper

1 Introduction

Artificial Intelligence is catalyzing an unprecedented structural reconfiguration of labour markets globally. Yet the regulatory frameworks governing employment relationships remain largely anchored to paradigms characteristic of twentieth-century industrial organization. This profound misalignment between technological innovation and regulatory adaptation constitutes a significant research gap, particularly regarding the identification and definition of the legal obligations incumbent upon employers in relation to workers' professional reskilling and skills development within contexts of AI-based automation processes and algorithmic management systems.

The scale of this challenge is formidable. The World Economic Forum estimates that AI and automation will lead to the loss of 85 million jobs globally by 2025, while simultaneously creating 97 million new roles. This transition generates a substantial governance challenge: workers displaced by traditional occupational roles must rapidly acquire new skills to participate in emerging labour markets. However, as emphasized in regulatory scholarship, no coherent legal framework currently allocates responsibility for such reskilling among employers, workers, and public institutions. This regulatory vacuum is particularly acute in the absence of international consensus on what constitutes adequate employer responsibility in technological transitions.

The research question guiding this study is therefore: **Which regulatory framework should govern employers' reskilling obligations when artificial intelligence replaces or substantially transforms occupational roles?** This question positions our analysis at the intersection of three critical domains: labour law, artificial intelligence governance, and organizational learning. We argue that sustainable responses to AI-driven labour market transformation require moving beyond traditional legal frameworks to incorporate insights from organizational learning theory, recognizing that reskilling represents not merely a compliance obligation but a fundamental transformation in how organizations acquire, develop, and deploy knowledge capital.

The significance of this inquiry extends beyond academic discourse. It addresses urgent practical concerns: How can employers fulfill their social responsibilities while maintaining competitive viability? What public support mechanisms are necessary to prevent labour market fragmentation? How can collective worker

participation shape technological transitions? These questions demand rigorous comparative analysis grounded in both legal doctrine and organizational theory.

2 Literature Review

2.1 Regulatory Dimensions of AI and Labour Markets

Existing literature examines the effects of Artificial Intelligence on labour markets from multiple disciplinary perspectives, with particular attention to regulatory implications and workers' rights. At the regulatory level, the European Union has taken a leading role through the adoption of the AI Act, although significant implementation inconsistencies and limited integration with labour law remain prominent concerns. From a comparative perspective, recent scholarship highlights substantial differences between Indonesia, the European Union, and the United States in the regulation of workforce governance, investment in human capital, and legal protection of workers affected by automation. Indonesian regulations remain fragmented, thereby hindering workforce reskilling, while the EU's more structured approach provides clearer guidance despite uneven implementation. The United States, notwithstanding its technological leadership, lacks federal AI workforce policies, resulting in regulatory fragmentation at the state level.

2.2 Algorithmic Management and Workers' Rights

With respect to algorithmic management and workers' rights, data-driven technologies tend to erode traditional employment protections, necessitating an integrated legislative adaptation through data protection, non-discrimination safeguards, and collective worker participation in digital human resource governance. Chen (2023) emphasizes that AI-based recruitment systems carry significant discriminatory risks, attributable to algorithmic bias stemming from unrepresentative datasets and distorted design choices. These insights underscore that reskilling obligations cannot be divorced from broader governance challenges around algorithmic fairness and worker dignity.

2.3 Reskilling, Skills Adaptation, and Industry 4.0

Regarding reskilling and skills adaptation, the implementation of Industry 4.0 generates widespread concerns related to job security, privacy, and technological trust, thereby requiring targeted interventions at both individual and organizational levels. From a broader perspective, a decent work framework grounded in job creation, social protection, workers' rights, and social dialogue emphasizes the need for preventive strategies capable of addressing both traditional and emerging risks associated with technological transformation.

Human resource management dimensions emerge prominently in analyses of organizational crisis-response strategies, which offer valuable insights for governing technological transitions. These studies stress the need for innovative solutions capable of balancing business continuity with workers' well-being and dignity during organizational change. This perspective is particularly relevant for understanding reskilling not as a reactive compliance measure but as a proactive organizational learning process.

2.4 Ethical and Human Rights Dimensions

The ethical dimensions of Artificial Intelligence are extensively explored in contemporary literature. Key risks associated with AI adoption include privacy violations, discrimination, unemployment, and security threats, necessitating comprehensive governance frameworks for ethically responsible implementation. Generative AI introduces specific employment-related risks including job losses, misinformation, data breaches, increased inequalities, and technostress, all of which significantly affect employers' responsibilities in technological adoption processes.

2.5 Global Justice Perspectives and AI Colonialism

From a global justice perspective, the concept of AI colonialism highlights how technological development exacerbates international inequalities through labour exploitation and environmental impacts in the Global South. This framework reveals that reskilling obligations cannot be understood solely through formal legal mechanisms but must account for structural inequalities in technological access and knowledge transfer between developed and developing economies.

2.6 Research Gap: Organizational Learning Theory Integration

Notably, existing literature examining AI and labour law has insufficiently integrated organizational learning theory, particularly the foundational scholarship on organizational learning processes (Nonaka and Takeuchi's knowledge creation theory, Argyris and Schön's double-loop learning, and Senge's systems thinking approach). This represents a critical gap, as reskilling constitutes fundamentally a process of organizational knowledge acquisition and dissemination. The absence of explicit engagement with organizational learning scholarship means that regulatory frameworks have largely overlooked how organizations systematically develop absorptive capacity for technological change. This gap is particularly significant for Track 25 of IFKAD (Organisational Learning in the Age of AI), which positions learning as central to sustainable organizational adaptation.

3 Research Framework and Methodology

3.1 Methodological Approach

This study adopts a comparative legal research methodology aimed at examining how different legal systems address the emerging issue of employers' obligations to reskill workers in contexts of artificial intelligence - driven technological transformation. Comparative legal analysis enables the systematic evaluation of regulatory approaches across jurisdictions, identifying both common principles and structural divergences in the governance of labour market transitions.

The research combines doctrinal legal analysis with insights from organizational learning theory. The doctrinal component involves the examination of primary and secondary legal sources - including legislation, regulations, case law, policy documents, and collective bargaining agreements - relevant to reskilling obligations and technological change in employment relationships. Organizational learning theory is incorporated as an analytical perspective to assess whether existing regulatory frameworks effectively support organizations' capacity to adapt to technological change through systematic knowledge development and workforce skill transformation. In this perspective, workforce reskilling is interpreted as a process of organizational knowledge creation and transfer, allowing the analysis to evaluate how legal frameworks influence firms' capacity to develop learning mechanisms and absorptive capacity in response to technological change.

The comparative analysis is conducted through three complementary analytical perspectives.

-First, normative analysis examines how existing legal instruments define, allocate, or fail to address employer responsibilities for workforce reskilling. This dimension focuses on the legal recognition of reskilling obligations within statutory labour law, regulatory policies, and judicial interpretation.

-Second, functional analysis evaluates how different legal systems pursue comparable regulatory objectives through diverse institutional mechanisms despite divergent formal legal structures. For example, collective bargaining agreements may perform regulatory functions similar to statutory provisions in different jurisdictions.

-Third, critical comparative analysis evaluates the effectiveness and coherence of these regulatory approaches by examining their capacity to support organizational adaptation to technological change. Drawing on organizational learning theory, this dimension assesses whether legal frameworks facilitate the development of institutional learning processes necessary for effective reskilling strategies.

3.2 Jurisdictional Selection Criteria

The comparative analysis focuses on four jurisdictional contexts selected according to three explicit criteria: regulatory relevance in the governance of artificial intelligence, diversity of labour law traditions and regulatory approaches, and representativeness of different levels of economic and institutional development.

Italy is included as a developed European jurisdiction characterized by a strong labour law tradition and an evolving regulatory approach to digital transformation. Its position within the European Union, while maintaining a distinct national legal system, provides insight into how supranational regulatory frameworks interact with domestic labour law institutions.

The European Union is examined as a supranational regulatory framework that has developed one of the most advanced legal approaches to AI governance, particularly through the adoption of the AI Act. The EU therefore represents a relevant case for assessing how emerging technological regulation interacts with existing labour law structures and how regulatory harmonization operates across diverse national contexts.

The United States is included as a global technological leader characterized by a decentralized and market-oriented regulatory model. The absence of a comprehensive federal framework addressing AI-related workforce

transitions offers a contrasting case to the more coordinated regulatory approach adopted within the European Union.

Finally, selected developing countries are considered to capture regulatory dynamics in contexts characterized by limited institutional capacity but increasing adoption of AI technologies. These jurisdictions highlight structural constraints affecting the implementation of reskilling policies and provide a broader global perspective on labour governance in technological transitions.

3.3 Analytical Framework

To ensure systematic comparison across jurisdictions, the study applies a five-dimensional analytical framework designed to assess the regulatory governance of employer reskilling obligations in the context of artificial intelligence.

From a knowledge management perspective, these dimensions also allow the evaluation of how regulatory frameworks shape organizational learning processes, particularly the capacity of firms to generate, transfer, and institutionalize new knowledge through workforce reskilling initiatives.

The first dimension, legal recognition, examines whether reskilling is explicitly recognized as an essential component of the employment relationship. This includes the presence of statutory obligations, collective bargaining provisions, or judicial interpretations establishing employer responsibilities for workforce skill development during technological transitions.

The second dimension, cost allocation, analyzes how the financial and organizational burdens of reskilling are distributed among employers, workers, and public institutions, revealing different regulatory philosophies regarding shared responsibility for workforce development.

The third dimension, algorithmic oversight, assesses institutional mechanisms for monitoring algorithmic management practices and their employment effects. These mechanisms may include regulatory supervision, algorithmic auditing procedures, transparency requirements, and institutional bodies responsible for oversight of automated decision-making systems affecting workers.

The fourth dimension, collective bargaining and worker participation, evaluates the role of collective worker participation in negotiating technological change and defining reskilling provisions. This dimension considers statutory frameworks enabling collective agreements as well as institutional mechanisms supporting social dialogue and dispute resolution.

The fifth dimension, corporate responsibility integration, examines how reskilling obligations are embedded within broader corporate governance and due diligence frameworks, including corporate social responsibility initiatives, supply chain accountability mechanisms, and organizational risk management practices.

Together, these five dimensions provide a structured analytical framework enabling systematic comparison of regulatory approaches across jurisdictions while identifying areas of regulatory convergence, divergence, and potential development. Each jurisdiction examined in the study is analyzed according to these five dimensions, allowing a structured comparison of regulatory approaches and highlighting both common patterns and jurisdiction-specific institutional responses to AI-driven labour market transformation.

4 Comparative Analysis of Regulatory Approaches

4.1 European Union Framework

The European Union represents the most developed regulatory response to AI governance through the AI Act. The AI Act establishes risk-based classification of AI systems and prescribes compliance obligations for high-risk applications, including those affecting employment relationships. However, the AI Act's integration with labour law remains underdeveloped. While the Act addresses algorithmic transparency and bias mitigation, it does not explicitly establish employer reskilling obligations.

Within EU member states, labour law traditions provide partially developed reskilling frameworks. Many EU jurisdictions recognize continuing education rights through collective bargaining or statutory frameworks, yet these provisions predate contemporary AI concerns and were not designed to address technological transformation at current scale. The challenge facing EU implementation is achieving coherence between AI governance and labour law modernization across 27 distinct member state systems while establishing enforceable mechanisms ensuring employers fulfill reskilling responsibilities.

4.2 Italian Labour Law Context

Italy exemplifies the challenge of regulatory adaptation within established labour law systems. Italian labour law contains sophisticated worker protection provisions, including continuing education rights within collective agreements. However, explicit legal recognition of reskilling obligations in contexts of AI-driven technological change remains absent from statutory frameworks. Italian jurisprudence has begun addressing algorithmic management issues, yet comprehensive frameworks addressing employer reskilling obligations during AI transitions have not yet crystallized in case law.

The Italian context reveals the insufficiency of traditional labour law categories for addressing contemporary technological challenges. Existing provisions on technological change, while protective, were formulated in industrial contexts and do not adequately address the acceleration and scale of AI-driven labour market disruption.

4.3 United States Fragmented Approach

The United States exemplifies fragmentation resulting from the absence of federal workforce AI policy. While OSHA (Occupational Safety and Health Administration) addresses worker safety, and the Equal Employment Opportunity Commission (EEOC) addresses algorithmic discrimination, no coordinated federal framework addresses employer reskilling obligations. This fragmentation produces varied outcomes across states. Some states have begun developing AI governance frameworks, yet labour-specific provisions remain underdeveloped.

The US market-based approach assumes that competitive labour markets incentivize employers to invest in reskilling without regulatory mandate. However, this assumption may be undermined by structural factors: employers may prioritize immediate cost reduction over long-term workforce development; information asymmetries prevent workers from assessing employer reskilling commitments; and labour market power imbalances limit worker capacity to negotiate reskilling provisions.

4.4 Developing Country Contexts

Developing countries face distinct constraints in establishing reskilling frameworks. Limited institutional capacity for regulatory enforcement, lower per-capita investment in workforce development, and rapid AI adoption imported from developed economies create contexts where reskilling obligations are theoretically recognized but practically difficult to implement. Global justice considerations become acute: developing countries experience AI colonialism dynamics where technological transformation originates from external sources without corresponding capacity-building support.

5 Organizational Learning Framework: Reconceptualizing Reskilling as Institutional Process

5.1 Organizational Learning Theory and Reskilling

A critical contribution of this paper is positioning reskilling within organizational learning theory frameworks. Reskilling should not be understood merely as compliance with employer obligations but as a fundamental process of organizational knowledge creation and capability development. This theoretical reframing opens important analytical possibilities.

According to Nonaka and Takeuchi's knowledge creation theory, organizations develop competitive advantage through systematic conversion of individual tacit knowledge into explicit organizational knowledge. In AI transition contexts, this framework illuminates how organizations must systematically capture and disseminate expertise regarding new technological systems. Reskilling obligations, from this perspective, represent institutionalized processes enabling this knowledge conversion.

Argyris and Schön's concept of double-loop learning is particularly relevant: organizations not only solve immediate problems (single-loop learning) but also examine and modify underlying assumptions (double-loop learning). Technological transitions demand double-loop learning—organizations must not merely update skills (single-loop) but fundamentally reconsider business processes, worker roles, and organizational purposes (double-loop).

Peter Senge's systems thinking approach emphasizes how organizations develop adaptive capacity through shared mental models, team learning, and systemic understanding. AI transitions demand precisely this systemic learning capacity: workers must understand how AI systems integrate with organizational workflows, how algorithmic decisions affect their work, and how to collectively adapt to technological changes.

5.2 Absorptive Capacity and Reskilling

Cohen and Levinthal's concept of absorptive capacity—the organization's ability to recognize, assimilate, and apply new external knowledge—provides crucial analytical tools for understanding reskilling effectiveness. Organizations with higher absorptive capacity more successfully integrate AI technologies because they possess requisite knowledge foundations and institutional mechanisms for technological learning. Reskilling obligations should thus be understood as investments in organizational absorptive capacity for technological change.

This perspective reveals why generic reskilling mandates may be insufficient: merely requiring training hours without ensuring organizations develop systematic capacities for knowledge integration produces compliance without effectiveness. Regulatory frameworks should incentivize organizational learning structures, not merely training provision.

5.3 Addressing the Organizational Learning Gap in Regulatory Frameworks

Most existing regulatory approaches to AI governance and labour law have not explicitly incorporated organizational learning theory. This represents a significant oversight. Regulatory frameworks emphasizing only compliance mechanisms (transparency requirements, bias audits) without requiring systematic organizational learning capacity development may achieve formal compliance while failing to ensure substantive worker protection or organizational adaptability.

The proposed framework therefore recommends that reskilling obligations explicitly require organizations to:

- Establish systematic mechanisms for identifying skills gaps created by AI implementation
- Develop institutional processes for converting tacit worker knowledge into explicit organizational learning
- Create feedback loops enabling workers to participate in continuous organizational adaptation to technological change
- Invest in collective learning mechanisms (communities of practice, cross-functional teams) addressing technological transitions
- Embed reskilling within organizational strategic planning rather than treating it as reactive compliance

6 Five-Dimensional Regulatory Framework: Detailed Analysis

6.1 Dimension 1: Legal Recognition of Reskilling as Employment Relationship Component

Current State: Most jurisdictions lack explicit legal recognition of reskilling as a fundamental employment component during AI-driven technological transitions. Even EU member states with strong worker protection traditions address reskilling through general provisions rather than AI-specific obligations.

Proposed Development: Legal frameworks should explicitly recognize reskilling as a core employer obligation during technological transitions, articulated through statutory provisions clarifying that AI implementation constitutes triggering events for reskilling rights. This recognition should extend beyond minimum training requirements to encompass employer responsibility for ensuring workers develop capability to perform successfully in transformed roles.

Implementation Mechanisms: National labour codes should be amended to define reskilling obligations in technology transition contexts, collective bargaining frameworks should include specific provisions enabling worker participation in defining reskilling needs, and case law development should establish judicial principles clarifying employer responsibilities.

6.2 Dimension 2: Cost-Sharing Mechanisms and Progressive Burden-Sharing

Current State: Cost allocation remains unclear across most jurisdictions, with ambiguity regarding whether reskilling constitutes an employer burden, worker responsibility, or public investment.

Proposed Development: A progressive burden-sharing model should allocate reskilling costs among employers (bearing primary responsibility for organizational transformation), workers (contributing time and learning effort), and public institutions (providing foundational support and ensuring equity across labour market participants). This allocation should reflect both fairness principles and incentive structures: employers implementing AI should bear primary costs, while public systems ensure that workers without employer reskilling access receive comparable support.

Implementation Mechanisms: Tax incentives could subsidize employer reskilling investments; public workforce development systems could be strengthened to provide foundational reskilling for displaced workers; and income support mechanisms could compensate workers during reskilling transitions.

6.3 Dimension 3: Independent Oversight of Algorithmic Management Practices

Current State: Oversight mechanisms addressing algorithmic management employment effects remain underdeveloped across most jurisdictions. While the EU AI Act addresses algorithmic transparency, labor-specific oversight remains limited.

Proposed Development: Independent oversight bodies should monitor algorithmic management practices, including recruitment algorithms, performance evaluation systems, and workplace surveillance technologies. These bodies should possess authority to audit algorithmic systems, receive worker complaints, investigate discriminatory outcomes, and enforce corrective measures.

Implementation Mechanisms: Specialized labour tribunals or ombudsperson offices could be established with algorithmic expertise; mandatory audits could be required for high-risk algorithmic systems affecting employment decisions; and transparency requirements could mandate disclosure of algorithmic decision-making criteria to affected workers and oversight authorities.

6.4 Dimension 4: Strengthening Collective Bargaining as Technological Change Governance Tool

Current State: Collective bargaining provides important protections in jurisdictions with strong labour movements (particularly EU) yet remains underdeveloped as a systematic framework for managing technological transitions. Most collective agreements address reskilling tangentially rather than positioning it centrally.

Proposed Development: Collective bargaining frameworks should be strengthened to enable workers to systematically participate in negotiating technological change, including: determining reskilling requirements, defining quality standards for training provision, establishing oversight mechanisms for algorithmic systems, and ensuring voice in workplace technology implementation.

Implementation Mechanisms: Labour law reforms could establish statutory rights for collective technology negotiation; bargaining structures could be reformed to enable sector-level standard-setting on technological transitions; and dispute resolution mechanisms could be strengthened to address collective disputes regarding reskilling obligations.

6.5 Dimension 5: Integration of Reskilling into Corporate Responsibility and Due Diligence Frameworks

Current State: Corporate social responsibility frameworks and due diligence requirements are increasingly addressing environmental and human rights dimensions but have insufficiently integrated labour transition responsibilities into AI implementation contexts.

Proposed Development: Due diligence standards should be expanded to require enterprises to assess potential reskilling needs before AI implementation, disclose expected labour market impacts, and establish remediation mechanisms when technological changes negatively affect workers. Supply chain responsibility should extend to ensuring reskilling support throughout value networks, not merely within headquarters organizations.

Implementation Mechanisms: Mandatory corporate impact assessments could be required before significant AI implementation; stakeholder engagement processes could ensure worker and community participation in technological transition planning; and remedy mechanisms could provide resources for workers harmed by inadequate reskilling support.

7 Expected Contributions and Practical Implications

7.1 Theoretical Contributions

This paper is expected to make several specific theoretical contributions based on the comparative analysis conducted across Italy, the European Union, the United States, and selected developing countries.

First, the analysis is expected to show that existing legal frameworks do not regulate employer reskilling obligations in a uniform or explicit manner but instead distribute responsibility through fragmented combinations of labour law, vocational training systems, collective bargaining arrangements, and public policy instruments. This finding contributes to labour law scholarship by clarifying that the central legal problem is not

merely the absence of regulation, but the lack of a coherent allocation of responsibility among employers, workers, and public institutions in contexts of AI-driven technological transition.

Second, the comparative analysis is likely to reveal that the most developed regulatory dimensions are not necessarily those directly addressing reskilling obligations. In particular, the European Union is expected to demonstrate stronger development in the areas of algorithmic oversight, transparency, and risk governance, especially through the AI Act, while the legal recognition of reskilling as a distinct employer obligation remains comparatively underdeveloped. This contributes to the literature by showing that advances in AI governance do not automatically translate into equally developed labour law protections concerning workforce transformation.

Third, the study is expected to demonstrate that different jurisdictional models reflect different regulatory philosophies. European systems are likely to rely more heavily on coordinated institutional arrangements involving public authorities, social dialogue, and collective bargaining, while the United States is expected to rely more on decentralized, market-oriented, and firm-level responses. Developing countries, by contrast, are likely to show that limited institutional capacity and uneven access to workforce development systems constrain the practical effectiveness of reskilling frameworks. This comparative insight contributes to interdisciplinary scholarship by identifying how legal institutions shape different models of adaptation to technological change.

Fourth, the paper contributes theoretically by integrating organizational learning and knowledge management perspectives into the legal analysis of AI governance. The analysis is expected to show that most current regulatory frameworks conceptualize reskilling primarily as a training or compliance issue, rather than as a process of organizational knowledge creation, knowledge transfer, and absorptive capacity building. This is an important theoretical contribution because it reframes reskilling as a mechanism through which organizations develop the internal capabilities required to adapt to technological transformation.

Fifth, the five-dimensional framework developed in this study is expected to provide a more precise comparative tool for future research. By distinguishing among legal recognition, cost allocation, algorithmic oversight, collective bargaining, and corporate responsibility integration, the framework makes it possible to identify not only whether a jurisdiction regulates technological transition, but also which dimensions are most developed, which remain weak, and how these dimensions interact. In this sense, the paper contributes both conceptually and methodologically to future comparative research at the intersection of labour law, AI governance, and organizational learning.

7.2 Practical Contributions for Policymakers

The expected findings of the comparative analysis also generate several practical implications for policymakers.

First, the study is likely to show that explicit legal recognition of employer reskilling obligations remains weak across jurisdictions, even where broader worker protection frameworks are relatively developed. This suggests that legislators should move beyond general training provisions and introduce more specific legal standards addressing workforce reskilling in contexts where AI significantly transforms or replaces occupational roles.

Second, the analysis is expected to reveal that cost allocation is one of the most unstable and underdefined regulatory dimensions. In most jurisdictions, it is likely that the burden of reskilling is not clearly distributed among employers, workers, and public institutions. This indicates the need for policymakers to design clearer burden-sharing mechanisms, including public support schemes, employer investment obligations, and protections for workers during retraining transitions.

Third, the comparative findings are expected to demonstrate that AI governance frameworks and labour law frameworks often remain insufficiently integrated. In the European context, for example, the AI Act may provide a relatively advanced structure for regulating algorithmic systems, but this does not necessarily ensure adequate labour protections related to reskilling. Policymakers should therefore strengthen the connection between technological regulation and employment law, ensuring that worker transition issues are addressed directly rather than indirectly.

Fourth, the analysis is likely to show that collective bargaining and social dialogue remain essential institutional mechanisms for governing technological change, especially in jurisdictions with stronger labour law traditions. This suggests that policymakers should reinforce legal frameworks that enable worker participation in decisions concerning AI adoption, job redesign, and reskilling strategies.

Fifth, for international organizations such as the ILO, OECD, and UN agencies, the expected findings indicate that developing countries may face particular structural barriers, including weaker institutional enforcement capacity, limited training infrastructure, and dependence on imported technologies. This suggests the importance of international standards and capacity-building initiatives aimed at supporting more equitable technological transitions across different economic contexts.

7.3 Implications for Business Practice

The paper is also expected to generate practical implications for business organizations adopting AI technologies in employment contexts.

First, the analysis is likely to show that organizations cannot treat reskilling as a purely voluntary or short-term human resource measure if they aim to manage technological transition sustainably. Instead, reskilling should be understood as a strategic organizational process linked to long-term workforce adaptation, institutional learning, and knowledge development.

Second, from an organizational learning perspective, the findings are expected to indicate that effective reskilling depends not only on the provision of training hours, but on the existence of internal mechanisms for identifying skills gaps, transferring knowledge, and integrating new competencies into organizational routines. This means that firms should invest in structured learning processes rather than relying on fragmented or reactive training interventions.

Third, the comparative analysis is likely to reveal that organizations operating in more coordinated institutional environments may benefit from stronger support structures, such as collective bargaining institutions, public training systems, and clearer regulatory guidance. By contrast, firms operating in more fragmented regulatory contexts may bear greater internal responsibility for designing reskilling strategies. This suggests that organizational responses to AI adoption are shaped not only by managerial choice but also by the institutional environment in which firms operate.

Fourth, the study is expected to show that firms introducing AI systems should strengthen their organizational absorptive capacity, namely their ability to recognize, assimilate, and apply new technological knowledge. In practice, this requires not only technical investments, but also the development of feedback loops, cross-functional learning mechanisms, and managerial processes that connect technological innovation with workforce capability development.

Finally, the findings suggest that integrating reskilling into broader frameworks of corporate responsibility and due diligence may become increasingly important for sustainable business governance. Organizations that fail to anticipate workforce transition needs may face not only legal and reputational risks, but also reduced capacity to adapt effectively to technological change. Conversely, firms that embed reskilling within organizational learning strategies are likely to be better positioned to combine innovation, workforce protection, and long-term competitiveness.

8 Conclusions and Future Research Directions

Artificial Intelligence is precipitating structural labour market transformations that existing regulatory frameworks were not designed to address. The absence of coherent legal frameworks governing employer reskilling obligations represents a critical regulatory gap with significant implications for worker welfare, organizational effectiveness, and social stability. This paper has demonstrated that sustainable governance of AI in labour markets requires:

First, explicit legal recognition of reskilling as a fundamental employer obligation during technological transitions, moving beyond optional corporate social responsibility toward binding regulatory requirements;

Second, the integration of organizational learning and knowledge management perspectives into regulatory frameworks, ensuring that reskilling is understood as a process of organizational knowledge creation, capability development, and absorptive capacity building rather than merely compliance-based training;

Third, development of cost-sharing mechanisms ensuring that reskilling investments are distributed equitably among employers, workers, and public institutions;

Fourth, strengthening of collective bargaining and worker participation mechanisms enabling workers to shape technological transitions rather than passively experiencing imposed change;

Fifth, integration of reskilling obligations into corporate due diligence frameworks, ensuring accountability throughout organizational value networks.

The comparative analysis reveals that while the European Union has advanced more developed regulatory frameworks, significant implementation inconsistencies persist and explicit labour law integration remains incomplete. The United States exhibits problematic fragmentation limiting worker protections. Developing countries face structural barriers requiring international support.

Future research should examine: implementation mechanisms for proposed frameworks across diverse jurisdictions; effectiveness assessment of different reskilling approaches utilizing organizational performance indicators; sectoral variations in reskilling needs and appropriate regulatory responses; and mechanisms ensuring global justice in technological transitions between developed and developing economies.

Ultimately, this paper argues that the social contract underlying employment relationships requires reconceptualization for the AI era. Technological innovation should develop in alignment with adequate worker protection and organizational learning capacity, ensuring that AI implementation contributes to inclusive prosperity rather than reproducing or amplifying labour market inequalities.

9 Ethics Declaration

This research did not require institutional ethical approval as it is a theoretical and comparative legal analysis based exclusively on published sources, policy documents, and legal instruments. No human subjects, sensitive personal data, or experimental procedures were involved.

10 AI Declaration

No AI tools were used in the conceptualization, research, or drafting of this paper. All analysis, comparative assessments, and written content were developed through manual research and writing processes.

References

Aloisi, A. (2024) 'Regulating Algorithmic Management at Work in the European Union: Data Protection, Non-discrimination and Collective Rights', *International Journal of Comparative Labour Law*, Vol. 2024, No. 1, pp. 1-25.

Chen, Z. (2023) 'Ethics and Discrimination in Artificial Intelligence-Enabled Recruitment Practices', *Palgrave Communications*, Vol. 9, No. 79.

De Stefano, V. (2022) 'The EU Proposal for an Artificial Intelligence Act and the Regulation of Work', *European Journal of Risk Regulation*, Vol. 13, No. 4, pp. 603-621.

Duong, C.D., Wach, K., Ejdys, J., Kazlauskaitė, R., Korzyski, P., Mazurek, G., Paliszkiewicz, J. and Ziemba, E. (2023) 'The Dark Side of Generative Artificial Intelligence: A Critical Analysis of Controversies and Risks of ChatGPT', *Entrepreneurship and Business Economics Review*, Vol. 11, No. 2, pp. 201-218.

Hamouche, S. (2021) 'Human Resource Management and the COVID-19 Crisis: Implications, Challenges, Opportunities, and Future Organizational Directions', *Journal of Management and Organization*, Vol. 27, No. 2, pp. 225-242.

Huang, C., Zhang, Z., Mao, B. and Yao, X. (2022) 'An Overview of Artificial Intelligence Ethics', *IEEE Transactions on Artificial Intelligence*, Vol. 3, No. 3, pp. 194-203.

Leesakul, N., Oostveen, A., Eimontaite, I., Wilson, M.L. and Hyde, R. (2022) 'Workplace 4.0: Exploring the Implications of Technology Adoption in Digital Manufacturing on a Sustainable Workforce', *Sustainability*, Vol. 14, No. 6, p. 3311.

Merola, R. (2022) 'Inclusive Growth in the Era of Automation and AI: How Can Taxation Help?', *Frontiers in Artificial Intelligence*, Vol. 5, p. 867832.

Regilme, S. (2024) 'Artificial Intelligence Colonialism: Environmental Damage, Labor Exploitation, and Human Rights Crises in the Global South', *South Asian Journal of International Studies*, Vol. 29, No. 3, pp. 950-958.

Schulte, P.A. et al. (2022) 'Occupational Safety and Health Staging Framework for Decent Work', *International Journal of Environmental Research and Public Health*, Vol. 19, No. 17, p. 10842.

Tjokro, G. et al. (2025) 'Reforming Legal Frameworks for Human Capital: Digital Strategies Driving Industry 5.0 and Sustainability', *Journal of Labour Law Review*, Vol. 6, No. 3, pp. 226-245.

Artificial Intelligence as Driver of Sustainable Performance in Entrepreneurial Ecosystems

Antonio Cimino[1], Vincenzo Corvello[1], Luca Famà[1], Ciro Troise[2]
[1]University of Messina, Messina, Italy
[2]University of Turin, Turin, Italy

Abstract

Entrepreneurial ecosystems are highly relevant in fostering startup success, as they facilitate access to essential resources and enhance collaboration among diverse actors through the promotion of knowledge and skills exchange. At the same time, the rapid growth in the use of artificial intelligence within organizations is not only profoundly transforming business practices but is also influencing the performance of startups within entrepreneurial ecosystems. Despite the increasing academic and managerial attention to these topics, the literature still presents an important gap regarding how startup behaviors, particularly their dynamic capabilities, influence the adoption of emerging and disruptive technologies such as artificial intelligence, and how such adoption can contribute to improving the sustainability performance of both startups and the entrepreneurial ecosystems in which they operate.

To address this gap, this study proposes a theoretical model that analyzes the role of startups' dynamic capabilities in the adoption of artificial intelligence. Furthermore, the study examines how the adoption of such technologies influences the sustainable performance of startups, considered a key indicator of the overall sustainability of entrepreneurial ecosystems. The model is empirically tested using Partial Least Squares Structural Equation Modeling (PLS-SEM), based on data collected through a self-administered questionnaire from startups listed in the official Register of Innovative Startups in Italy.

The results show that sensing and transforming capabilities positively influence the adoption of artificial intelligence technologies, while seizing capabilities do not exhibit a significant effect. Moreover, it is found that the adoption of artificial intelligence has a positive impact on the sustainability of entrepreneurial ecosystems. These results contribute to enriching the literature on entrepreneurial ecosystems and provide relevant implications for both researchers and policymakers, as well as managers interested in promoting sustainable innovation.

In particular, they highlight the need to strengthen startups' capabilities for adaptation and technological monitoring, as well as to develop targeted support policies that facilitate access to financial resources, managerial skills, and training programs, in order to foster more effective adoption of artificial intelligence and, consequently, improve the sustainability and long-term resilience of entrepreneurial ecosystems.

Keywords: Dynamic Capabilities, Artificial Intelligence, Entrepreneurial Ecosystems, Startups, Sustainable Performance

Paper type: Academic Research Paper

1 Introduction

Entrepreneurial Ecosystems (EEs) provide access to essential resources and improve the collaboration capabilities of startups, thus also facilitating knowledge exchange. Evidence highlights that startups operating within well-structured EEs are better prepared to face innovation-related challenges such as the limited availability of financial resources and restricted access to knowledge (Franco-Leal and Diaz-Carrion, 2020). Although, as stated, these ecosystems play a fundamental role in the success of startups, it is equally important how startups position themselves within them, since the startups behavior can significantly influence the overall effectiveness of the ecosystem. In the literature, the dynamic capabilities framework is often used precisely to analyze startup behavior. When referring to dynamic capabilities, it means the set of organizational competencies that enable firms to adapt, respond effectively, and succeed in rapidly changing business environments (Teece, 2014). At the same time, the success of an ecosystem has also been linked to the sustainability performance of its members (Cobben et al., 2023). If EEs exist and are particularly relevant for the success of startups, the dynamic capabilities of their members and their sustainability performance are therefore very important. For this reason, in recent years researchers have paid increasing attention to this relationship. Russo (2003) used the dynamic capabilities model to analyze how the introduction of new process standards can strengthen the ability of manufacturing facilities to improve environmental performance, thus reducing harmful emissions. Chakrabarty and Wang (2012) adopted this perspective to analyze how firms develop and implement sustainable practices, while Bhupendra and Sangle (2015) provided significant contributions by highlighting the role of dynamic capabilities in pollution prevention strategies. Mousavi et al. (2018) through a survey-based study, demonstrated that the main dimensions of dynamic capabilities have a positive impact on

sustainability-driven innovation. To observe the current scenario comprehensively, it is also necessary to consider that in recent years an increasing number of disruptive technologies have been emerging, which are profoundly reshaping organizations and consequently influencing business performance. Among these technologies, Artificial Intelligence (AI) stands out as one of the most disruptive (Arsenyan and Piepenbrink, 2023; Zhang et al., 2024), driving technological change across different sectors (Almansour, 2023). In this research scenario, a gap emerges regarding the specific study of how startup behavior influences the integration of disruptive technologies into their processes to improve the sustainability of their EEs. To fill this gap, this study proposes a theoretical model that examines how startups' dynamic capabilities influence the adoption of artificial intelligence and how the latter can influence their sustainable performance, a key element for improving the sustainability of EEs.

2 Theory

2.1 Theoretical Model

This study proposes a theoretical model that investigates how start-ups' dynamic capabilities influence AI adoption (AIA) and how this latter can impact startups sustainability performance (SP) and in turn their EE sustainability. Startups' behavior is operationalized through the three core components of the dynamic capabilities framework: sensing (SEN), seizing (SEI), and transforming (TRA) (Teece, 2014). AIA refers to the degree to which startups incorporate AI technologies into their operations and business activities (Almaiah et al., 2022; Chen and Tajdini, 2024), while SP refers to a firm's commitment to environmental and social responsibility in terms of resource efficiency, environmental protection, and long-term sustainability in its operations and business model (Shen et al., 2020; Rauter et al., 2018). The proposed theoretical framework is illustrated in Figure 1.

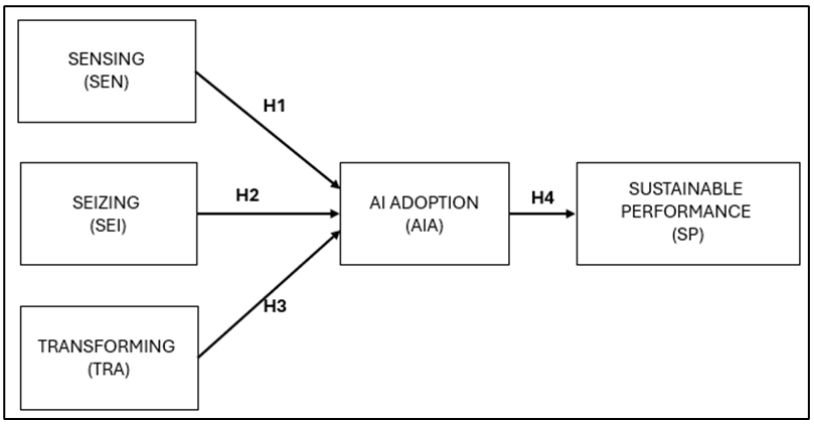


Figure 1: Theoretical model

2.1 Hypothesis development

Dynamic capabilities and Artificial Intelligence adoption

Dynamic capabilities refer to the firm's abilities to adapt, evolve, and maintain competitiveness in dynamic and uncertain environments (Corvello et al., 2023). Teece (2007, 2014) proposes a dynamic capabilities framework articulated around three interrelated components: sensing, seizing and (re)configuring. Sensing implies the ability of a firm to observe the external environment, intercepting and potentially anticipating imminent changes, thus enabling the integration of, for example, new emerging technologies and the capture of new market opportunities (Teece, 2014). Seizing, on the other hand, is the ability to effectively capitalize on the opportunities identified through sensing by allocating resources and developing timely strategies that respond to new opportunities (Teece, 2014). (Re)configuring, referred to as transforming, represents the organization's operational capabilities to realign and reconfigure internal resources, processes, and structures to support all the changes stimulated by the first two components of the framework (Teece, 2014). The literature has consistently highlighted how dynamic capabilities represent an important tool in determining whether and to what extent firms are ready to adopt new technologies (Graham and Moore, 2021; Majhi et al., 2023; Aljarboa, 2024). Based on this literature, we propose that the development of dynamic capabilities significantly improves startups' ability to adopt and implement disruptive technologies such as AI.

H1: Sensing (SEN) positively affects Artificial Intelligence Adoption (AIA) by startups

H2: Seizing (SEI) positively affects Artificial Intelligence Adoption (AIA) by startups

H3: Transforming (TRA) positively affects Artificial Intelligence Adoption (AIA) by startups

Artificial Intelligence Adoption and Sustainable performance

AI is considered one of the most revolutionary and impactful technologies of recent years, as it is radically changing the way companies manage all activities, thus improving performance across a wide range of industries (Arsenyan and Piepenbrink, 2023; Zhang et al., 2024). For the purposes of this study, AIA refers to the degree to which startups integrate AI technologies into their operations and into value-generating processes (Almaiah et al., 2022; Chen and Tajdini, 2024). An increasing number of studies are focusing on the various organizational implications deriving from AIA, highlighting positive results across multiple performance dimensions (Nayal et al., 2021; Belhadi et al., 2021; Abrokwah-Larbi and Awuku-Larbi, 2023; Necula et al., 2024). Building on these results, we therefore propose that the adoption of AI technologies by startups can play a key role in also improving sustainability performance within EEs.

H4: Artificial Intelligence Adoption (AIA) positively influence startups sustainable performances (SP)

3. Methodology

3.1 Measurement instrument

To achieve the research objective of this study, a structured questionnaire-based survey was conducted. Measurement items for each construct were drawn from established scales in the existing literature and, where appropriate, were slightly adapted to fit the specific context of this study. The final questionnaire consisted of 21 items, each measured using a 7-point Likert scale ranging from “strongly disagree” to “strongly agree”. A detailed overview of all scale items is presented in Table 1.

Table 1: Variables and measures

Variable	Items	Measurement items	Source
Sensing (SEN)	SEN1	Our company knows the best practices in the market	Kump and Kessler (2019)
	SEN2	Our company is up to date on the current market situation	
	SEN3	Our company systematically searches for information on the current market situation	
	SEN4	As a company, we know how to access new information	
Seizing (SEI)	SEI1	Our company can quickly relate to new knowledge from the outside	Kump and Kessler (2019)
	SEI2	We recognize what new information can be utilized in our company	
	SEI3	Our company is capable of turning new technological knowledge into process and product innovation	
Transforming (TRA)	TRA1	By defining clear responsibilities, we successfully implement plans for changes in our company	Kump and Kessler (2019)
	TRA2	Even when unforeseen interruptions occur, change projects are seen through consistently in our company	
	TRA3	Decisions on planned changes are pursued consistently in our company	
	TRA4	In the past, we have demonstrated our strengths in implementing changes	
	TRA5	In our company, change projects can be put into practice alongside the daily business	
AI Adoption (AIA)	AIA1	My company is ready to adopt AI technology to generate innovations	Adapted from Almaiah et al. (2022) and from Chen and Tajdini (2024)
	AIA2	My company is willing to upgrade their technology and ready to include AI as part of it	
	AIA3	We have implemented AI in all our company processes	
	AIA4	Relative to the potential of AI for our company activities, our AI implementation is extensive	
	AIA5	AI has substantially changed our business processes	
Sustainable performance (SP)	SP1	Our company has always attached great importance to reducing environmental pollution	Shen et al. (2020), Rauter et al. (2018)
	SP2	Our company gives priority on resource efficiency in production	
	SP3	Our company puts great emphasis on social responsibility	
	SP4	Our company regards product sustainability as its top priority	
	SP5	Our company is focused on building a sustainable business model	

3.2 Data collection and data analysis

The survey was administered via email between November 15, 2024, and January 10, 2025, addressing startups registered in the official Italian Innovative Startup Business Register. Overall, 234 fully completed responses were collected. However, 24 responses contained missing data and were excluded from further analysis, resulting in a final sample of 210 valid responses. This dataset was used to empirically test the proposed research model and hypotheses. Partial Least Squares Structural Equation Modeling was employed for the analysis, given its suitability for studies with relatively small sample sizes (Hair et al., 2021) and its effectiveness in exploratory research settings (Hair et al., 2017).

4 Results

4.1 Descriptive results

Table 2 presents the demographic and organizational characteristics of the startups included in the final sample (n = 210). The majority of the startups operate within the service activities and utilities sector, which accounts for 58% of the sample. This is followed by industry and manufacturing (22%) and ICT (10%). Other sectors such as R&D, crafts, tourism, and healthcare are marginally represented, each comprising 4% or less of the sample. This distribution reflects the prominence of service-based and technology-driven ventures within Italy's innovative startup landscape. In terms of regional distribution, half of the surveyed startups (50%) are located in Northern Italy, followed by 27% in the Central regions and 23% in the South and Islands. This regional concentration aligns with broader economic trends, where Northern Italy traditionally hosts a higher density of entrepreneurial activity and innovation infrastructure. Finally, most startups in the sample are small-sized enterprises: 81% have 10 or fewer employees. Only 11% report having between 11 and 20 employees, while 8% employ more than 20 people. This indicates that the sampled startups are predominantly micro-enterprises, consistent with the early-stage nature of ventures typically found in innovation registries.

Table 2: Characteristics of the sample

Characteristics		Frequency (n)	Percentage (%)
Industry sector	Service activities and utilities	122	58%
	Industry and Manufacturing	45	22%
	ICT	21	10%
	R&D	9	4%
	Crafts	9	4%
	Tourism	2	1%
	Tourism	2	1%
Geographical location	North	105	50%
	Centre	57	27%
	South and Islands	48	23%
Number of employees	≤10	170	81%
	11-20	23	11%
	≥21	17	8%

4.2 Measurement model

To evaluate the measurement model, indicator reliability, internal consistency validity, convergent validity and discriminant validity were conducted. Indicator reliability was confirmed through the analysis of indicator loadings, all of which surpassed the recommended threshold of 0.70, indicating satisfactory individual item reliability (Hair et al., 2021). Internal consistency reliability was assessed using composite reliability (Rho_c), with all constructs exceeding the acceptable benchmark values, thus demonstrating strong internal coherence among the indicators. Convergent validity was evaluated via the Average Variance Extracted (AVE), with all constructs showing AVE values above the 0.50 threshold, confirming that the indicators adequately capture the underlying constructs (Hair et al., 2021). Detailed results for reliability and convergent validity are presented in Table 3. In addition, discriminant validity was assessed using the Fornell-Larcker criterion (Fornell and Larcker, 1981). The analysis confirmed that, for each construct, the square root of the AVE was greater than its correlations with any other construct, supporting the distinctiveness of each latent variable. These results are shown in Table 4 and provide strong evidence for the discriminant validity of the measurement model.

Table 3: Internal consistency and convergent validity

Construct	Rho_C	AVE
SEN	0.869	0.626
SEI	0.854	0.661
TRA	0.912	0.676
AIA	0.936	0.745
SP	0.766	0.625

Table 4: Discriminant validity

Fornell-Larcker					
	SEN	SEI	TRA	AIA	SP
SEN	0.791				
SEI	0.65	0.813			
TRA	0.542	0.635	0.822		
AIA	0.323	0.364	0.472	0.863	
SP	0.468	0.492	0.638	0.733	0.79

4.3 Structural model and hypothesis testing

The results of the hypothesis testing are summarized in Table 5, which reports the path coefficients representing the estimated relationships within the structural model. The table also includes key statistical outputs such as standardized beta (β) values, sample means, standard deviations, t-values, and confidence intervals. Complementing this, Figure 2 visually presents these relationships alongside the R^2 values, providing an overview of the model's explanatory power. The results confirm the predictive validity of the model. Analysis of the structural paths reveals that SEN and TRA capabilities both have a significant positive impact on AIA, thereby supporting hypotheses H1 and H3. In contrast, SEI capability does not show a statistically significant effect on AIA, resulting in the rejection of hypothesis H2. Furthermore, AIA demonstrates a positive and significant relationship with SP, thereby confirming hypothesis H4.

Table 5: Hypothesis testing

Path coefficient	Orig. Mean	Bootstrap Mean	Bootstrap SD	T Stat	2.5% CI	97.5% CI	Result
SEN à AIA	0.144	0.15	0.067	2.142	0.015	0.279	Yes
SEI à AIA	0.084	0.088	0.094	0.901	-0.105	0.263	No
TRA à AIA	0.33	0.333	0.087	3.796	0.161	0.503	Yes
AIA à SP	0.693	0.699	0.045	15.25	0.606	0.785	Yes

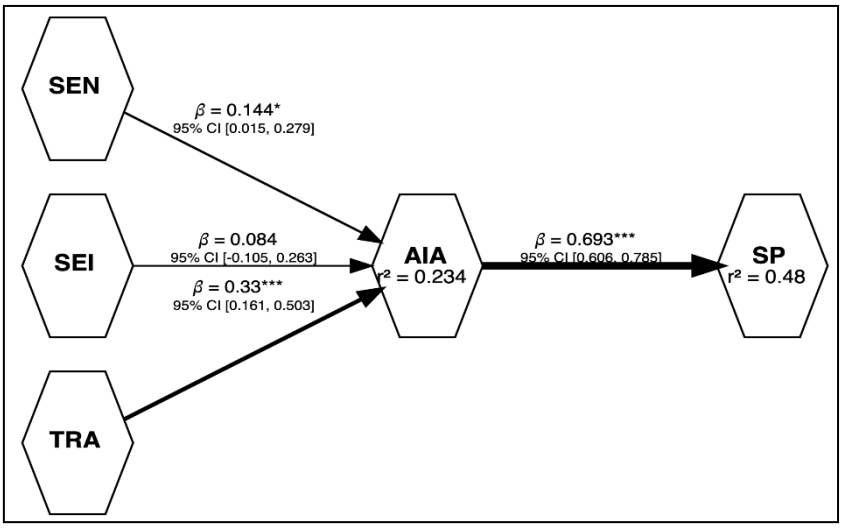


Figure 2: Results of the final model evaluation

5 Discussion

The first set of results highlights how SEN and TRA capabilities have a positive and significant influence on AIA, thus confirming hypotheses H1 and H3. These results therefore confirm that innovative startups, which

operate in dynamic and highly technology-intensive contexts and master the ability to continuously monitor the external environment (sensing) and reconfigure internal processes (transforming), are able to exploit the potential of emerging technologies such as AI. The findings of this study therefore confirm the existing literature that emphasizes the importance of sensing and transforming dynamic capabilities in making firms agile and responsive to technological progress, facilitating the integration of tools such as AI (Majhi et al., 2023; Graham and Moore, 2021; Aljarboa, 2024).

As for the second set of results, it appears that SEI does not have a relevant effect on AIA, thus not supporting hypothesis H2 and contrasting several previous studies that also identified seizing as a key driver of technological adoption (Majhi et al., 2023; Graham and Moore, 2021; Aljarboa, 2024). This misalignment with the results of the existing literature can be explained by the type of companies considered; since these are innovative startups, this type of firms, still in the early stages of development, are able to recognize AI-related opportunities (sensing) and to fairly easily transform their structures and processes (transforming), but encounter practical difficulties such as financial constraints and limited managerial experience, which hinder their ability to seize and effectively implement these opportunities. In other words, while they may have the intent and strategic vision, their execution capability may not be sufficiently mature to translate opportunities into concrete AI adoption. This contextual specificity reflects the structural and operational challenges that innovative startups often face, especially in their early stages.

The study also finds that AIA has a positive effect on startups' SP, thus confirming hypothesis H4. This result supports the growing body of literature indicating that AI technologies can significantly improve firm performance across multiple dimensions, including sustainability outcomes (Nayal et al., 2021; Belhadi et al., 2021; Abrokwah-Larbi and Awuku-Larbi, 2023; Necula et al., 2024). Within the context of entrepreneurial ecosystems, the sustainable performance of individual startups contributes to the overall sustainability of the ecosystem itself (Cobben et al., 2023). As startups adopt AI to improve their environmental, social, and economic outcomes, they collectively strengthen the resilience and long-term sustainability of the ecosystems in which they operate.

6 Implications

The results of this study offer several important practical applications useful for startup founders, innovation policymakers and all stakeholders involved in entrepreneurial ecosystems. First of all, given the importance of sensing and transforming capabilities on AI adoption, all startups should invest in the development of strategic agility and internal adaptability in order to fully exploit new technological opportunities. This goal can be achieved through the creation of mechanisms to continuously monitor technological trends and flexibly reconfigure their processes. Secondly, the absence of a significant relationship between seizing capabilities and AI adoption suggests that early-stage startups may face barriers in execution, such as limited financial or managerial resources. This highlights the importance of targeted support measures, such as access to funding, mentoring, and training, that can help startups move from opportunity recognition to effective implementation. As for the positive impact of AI adoption on sustainable performance, fostering AI readiness within startups can represent a lever to improve the competitiveness of individual firms as well as to advance the sustainability of the broader entrepreneurial ecosystem. Policymakers and ecosystem developers should therefore consider integrating support for digital transformation and sustainability into innovation programs to ensure the long-term resilience of ecosystems.

7 Conclusions

This study helps to better understand how startups' dynamic capabilities influence the adoption of artificial intelligence and how this contributes to improving their sustainable performance, seen as an expression of the sustainability of the entire entrepreneurial ecosystem. Drawing on the dynamic capabilities framework, a theoretical model was developed and then empirically tested through data collected via a survey from innovative startups registered with the Italian Chamber of Commerce. The results show that sensing and transforming capabilities significantly foster AI adoption, with positive effects on startups' sustainable performance. An interesting aspect concerns the seizing capability, which does not show a significant effect on AI adoption, thus offering a more articulated view of the difficulties that startups, especially in their early stages and with limited resources, face when implementing disruptive technologies. Despite the results obtained, the study presents some limitations. First, the use of cross-sectional data, collected at a single point in time, does not allow observation of how the relationships between dynamic capabilities and AI adoption evolve over time. Longitudinal studies could provide a more complete and dynamic view of these dynamics in the long term. Moreover, the sample analyzed is composed exclusively of Italian innovative startups, which may limit the

possibility of generalizing the results to other geographic or sectoral contexts. For this reason, future research could extend the analysis by including startups from other countries or with different levels of maturity, in order to test the robustness of the model in different contexts. Finally, qualitative studies could further explore why, in some cases, seizing capabilities are less effective, offering a deeper understanding of the difficulties related to AI implementation.

7 Ethics Declaration

The authors declare that this study did not require formal ethical approval.

8 AI Declaration

The authors declare that generative AI tools were used exclusively for language proofreading.

References

Abrokwah-Larbi, K. and Awuku-Larbi, Y. (2023) "The impact of artificial intelligence in marketing on the performance of business organizations: evidence from SMEs in an emerging economy", *Journal of Entrepreneurship* in Emerging Economies, Vol. 16, No. 4, https://doi.org/10.1108/JEEE-07-2022-0207

Aljarboa, S. (2024) "Factors influencing the adoption of artificial intelligence in e-commerce by small and medium-sized enterprises", *International Journal of Information Management Data Insights*, Vol. 4, 100285, https://doi.org/10.1016/j.jjimei.2024.100285

Almaiah, M.A. et al. (2022) "Measuring Institutions' Adoption of Artificial Intelligence Applications in Online Learning Environments: Integrating the Innovation Diffusion Theory with Technology Adoption Rate", *Electronics*, Vol. 11, 3291, https://doi.org/10.3390/electronics11203291

Almansour, M. (2023) "Artificial intelligence and resource optimization: A study of Fintech start-ups", *Resources Policy*, Vol. 80, 103250, https://doi.org/10.1016/j.resourpol.2022.103250

Arsenyan, J. and Piepenbrink, A. (2023) "Artificial intelligence research in management: a computational literature review", *IEEE Transactions on Engineering Management*, Vol. 71, pp 5088–5100, https://doi.org/10.1109/TEM.2022.3229821

Belhadi, A. et al. (2021) "Artificial intelligence-driven innovation for enhancing supply chain resilience and performance under the effect of supply chain dynamism: an empirical investigation", *Annals of Operations Research*, Vol. 333, No. 2, pp 627–652, https://doi.org/10.1007/s10479-021-03956-x

Bhupendra, K.V. and Sangle, S. (2015) "What drives successful implementation of pollution prevention and cleaner technology strategy? the role of innovative capability", *Journal of Environmental Management*, Vol. 155, pp 184–192, http://dx.doi.org/10.1016/j.jenvman.2015.03.032

Chakrabarty, S. and Wang, L. (2012) "The long-term sustenance of sustainability practices in MNCs: a dynamic capabilities perspective of the role of R&D and internationalization", *Journal of Business Ethics*, Vol. 110, pp 205–217, http://dx.doi.org/10.1007/s10551-0121422-3

Chen, J. and Tajdini, S. (2024) "A moderated model of artificial intelligence adoption in firms and its effects on their performance", Information Technology and Management, https://doi.org/10.1007/s10799-024-00422-5

Cobben, D., Ooms, W. and Roijakkers, N. (2023) "Indicators for innovation ecosystem health: A Delphi study", *Journal of Business Research*, Vol. 162, 113860, https://doi.org/10.1016/j.jbusres.2023.113860

Corvello, V., Cimino, A. and Felicetti, A.M. (2023) "Building start-up acceleration capability: A dynamic capability framework for collaboration with start-ups", *Journal of Open Innovation: Technology, Market, and Complexity*, Vol. 9, 100104, https://doi.org/10.1016/j.joitmc.2023.100104

Fornell, C. and Larcker, D.F. (1981) "Evaluating structural equation models with unobservable variables and measurement error", *Journal of Marketing Research*, Vol. 18, No. 1, pp 39–50

Franco-Leal, N. and Diaz-Carrion, R. (2020) "A dynamic analysis of the role of entrepreneurial ecosystems in reducing innovation obstacles for startups", *Journal of Business Venturing Insights*, Vol. 14, e00192, https://doi.org/10.1016/j.jbvi.2020.e00192

Graham, K.W. and Moore, R.S. (2021) "The Role of Dynamic Capabilities in Firm-Level Technology Adoption Processes: A Qualitative Investigation", *Journal of Innovation Management*, Vol. 9, No. 1, pp 25–50, http://dx.doi.org/10.24840/2183-0606_009.001_0004

Hair, J.F. et al. (2017) Advanced Issues in Partial Least Squares Structural Equation Modeling, Sage Publications

Hair, J.F. et al. (2019) "When to use and how to report the results of PLS-SEM", *European Business Review*, Vol. 31, No. 1, pp 2–24, https://doi.org/10.1108/EBR-11-2018-0203

Hair, J.F. et al. (2021) Partial Least Squares Structural Equation Modeling (PLS-SEM) Using R, Springer, Cham

Kump, B. and Kessler, A. (2019) "Toward a dynamic capabilities scale: measuring organizational sensing, seizing, and transforming capacities", *Industrial and Corporate Change*, Vol. 28, No. 5, pp 1149–1172, http://dx.doi.org/10.1093/icc/dty054

Majhi, S.G., Mukherjee, A. and Anand, A. (2023) "Business value of cognitive analytics technology: a dynamic capabilities perspective", *Journal of Information and Knowledge Management*, Vol. 53, No. 6, pp 1231–1249, https://doi.org/10.1108/VJIKMS-07-2021-0128

Mousavi, S., Bossink, B. and van Vliet, M. (2018) “Dynamic capabilities and organizational routines for managing innovation towards sustainability”, *Journal of Cleaner Production*, Vol. 203, pp 224–239, https://doi.org/10.1016/j.jclepro.2018.08.215

Nayal, K. et al. (2021) “Exploring the role of artificial intelligence in managing agricultural supply chain risk to counter the impacts of the COVID-19 pandemic”, *International Journal of Logistics Management*, Vol. 33, No. 3, pp 744–772, https://doi.org/10.1108/IJLM-12-2020-0493

Necula, S.C., Fotache, D. and Rieder, E. (2024) “Assessing the impact of artificial intelligence tools on employee productivity: insights from a comprehensive survey analysis”, *Electronics*, Vol. 13, No. 18, 3758, https://doi.org/10.3390/electronics13183758

Rauter, R., Globocnik, D., Perl-Vorbach, E. and Baumgartner, R.J. (2018) “Open innovation and its effects on economic and sustainability innovation performance”, *Journal of Innovation and Knowledge*, Vol. 4, pp 226–233, https://doi.org/10.1016/j.jik.2018.03.004

Russo, M.V. (2003) “The emergence of sustainable industries: building natural capital”, *Strategic Management Journal*, Vol. 24, No. 4, pp 317–331, https://doi.org/10.1002/smj.298

Shen, J., Sha, Z. and Wu, Y.J. (2020) “Enterprise Adaptive Marketing Capabilities and Sustainable Innovation Performance: An Opportunity-Resource Integration Perspective”, *Sustainability*, Vol. 12, 469, https://doi.org/10.3390/su12020469

Teece, D.J. (2007) “Explicating dynamic capabilities: The nature and microfoundations of (sustainable) enterprise performance”, *Strategic Management Journal*, Vol. 28, No. 13, pp 1319–1350, https://doi.org/10.1002/smj.640

Teece, D.J. (2014) “The foundations of enterprise performance: dynamic and ordinary capabilities in an (economic) theory of firms”, *Academy of Management Perspectives*, Vol. 28, No. 4, pp 328–352, https://doi.org/10.5465/amp.2013.0116

Willaby, H.W. et al. (2015) “Testing complex models with small sample sizes: a historical overview and empirical demonstration of what partial least squares (PLS) can offer differential psychology”, *Personality and Individual Differences*, Vol. 84, pp 73–78, https://doi.org/10.1016/j.paid.2014.09.008

Zhang, X., Yu, P. and Ma, L. (2024) “How and when generative AI use affects employee incremental and radical creativity: an empirical study in China”, *European Journal of Innovation Management*, in press, https://doi.org/10.1108/EJIM-04-2024-0466

Innovation and Sustainability as Judged by the Market

Maria Teresa Bianchi, Raffaele de Socio, Federica Mastrantonio
Sapienza University of Rome, Rome, Italy

Abstract

The aim of this research is to describe the impact of innovative and sustainable investments made by listed companies on their market performance. The research methodology applied to this paper is the qualitative method through the application of the multiple case study analysis. Sources of analysis were company and market data, financial statements and sustainability reports for the financial years 2022, 2023 and 2024 of a sample of Italian listed companies selected for their commitment to innovative and sustainable investments. We investigated the relationship between these investments and their market capitalisation. The results achieved demonstrate a positive relationship between investments and market performance. It emerged that companies which incorporate such investments into their strategies, drawing on their knowledge management expertise, make the market aware of their operational efficiency (as well as creating shared value) and attract investors willing to invest in such companies. In fact, the market views the company's investment commitment favourably, as it is seen as the foundation for greater operational efficiency. Innovative and sustainable investments lead to higher market prices, thereby contributing to growth in market capitalisation. The study focuses not only on the promotion and execution of investments by listed companies, but also on their ability and willingness to both innovate and operate sustainably, by investing substantial resources towards these ends, and on the market's response to such actions. This represents a departure from the traditional literature, which focuses primarily on a company's ability to invest in either innovation or sustainability but has not yet explored the impact that joint investment in both areas may have on market capitalisation.

The paper enriches the existing literature through studies and findings that provide a useful reference point for both the academic and practical implications of developments in research on the adoption of innovative and sustainable corporate investments, and the functional link between these and market reactions.

Keywords: Innovation; Sustainability; ESG; Market capitalisation; Operational efficiency.

Paper type: Practical Paper.

1 Introduction

The study is focused on the impact of innovative and sustainable investments made by listed companies on their market performance.

Today, businesses are expected to demonstrate their value to potential investors and the market, as well as their ability to generate that value.

Including investments aimed at fostering innovation and sustainability in strategic or business plans is a winning strategy for demonstrating operational efficiency to the market and attracting new investors.

The purpose of this research is to understand how the market values innovative and sustainable companies, including in terms of market capitalisation. The research methodology applied to this paper is the qualitative method through the application of the multiple case study analysis. Sources of analysis were company and market data, financial statements and sustainability reports for the financial years 2022, 2023 and 2024 of a sample of Italian listed companies selected for their commitment to innovative and sustainable investments. We investigated the relationship between these investments and their market capitalisation.

We have seen how companies that are most committed to innovative and sustainable investment are valued by investors and rewarded by the market. The research focuses not only on the promotion and execution of investments by listed companies, but also on their ability and willingness to both innovate and operate sustainably, by investing substantial resources towards these ends, and on the market's response to such actions. This represents a departure from the traditional literature, which focuses primarily on a company's ability to invest in either innovation or sustainability but has not yet explored the impact that joint investment in both areas may have on market capitalisation.

The paper enriches the existing literature through studies and findings that provide a useful reference point for both the academic and practical implications of developments in research on the adoption of innovative and sustainable corporate investments, and the functional link between these and market reactions.

The remainder of this paper is structured as follows: section 2 provides a literature review that includes our research question; section 3 shows research methods and multiple case study; section 4 explains our results and discussion; section 5 outlines our conclusions; section 6 and 7 express ethics declaration and AI declaration.

2 Literature Review

The integration of ESG investments remains essential for driving ethical, environmentally sustainable and socially responsible innovation. It is very important, in fact, that the theme of innovation is also combined with social aspects and performance (Deng et al., 2025).

The resource-based view (RBV) argues that companies achieve superior performance by exploiting resources and capabilities that are valuable, rare, inimitable and non-substitutable (Wernerfelt, 1984; Barney, 1991).

In this context, therefore, the implementation of ESG principles does not simply act as an isolated resource, but as a complex system of capabilities developed through innovative investments, supported by environmental processes, human capital and governance routines. Accordingly, investors in these sectors are willing to pay for the risk of innovation in the hope of obtaining higher future returns (Fatemi et al., 2018). It is therefore necessary to define the exact contours of innovation and, above all, to understand how it is interpreted by the entrepreneur – or rather by the entrepreneurship factor (C. Bianchi, 1998; P. Capaldo, 2013) – and by investors (Samy ElDeeb et al., 2025). Bearing in mind that truly innovative investments are directed and channelled towards improving the social pillar (C. Bianchi, 2010).

Recently, companies have focused on the extent to which sustainable investments can increase market value, given their proportional correlation with market capitalisation (Raimo et al., 2020; Aydogmus et al., 2022; Zhou et al., 2022; Ni and Sun, 2023; Espinosa-Mendez et al., 2024).

Although still in its early stages, the integration of ESG measures appears to have the potential to influence a company's financial value and performance (Nirino et al., 2021; Alkaraan et al., 2022; García-Amate et al., 2023) and market value (Lo and Kwan, 2017; Janicka and Sajnog, 2022; Zhou et al., 2022). In fact, with the integration of innovative and sustainable strategies and policies into portfolio theory, ESG could potentially become an essential factor influencing investment in company shares (Li et al., 2022).

Environmental investments reduce energy consumption and waste, thereby lowering operating costs and increasing resource productivity (Wong et al., 2020; Vagnoni et al., 2025). Social investments, including comprehensive employee training and safety systems, strengthen human capital and reduce staff turnover, resulting in fewer operational disruptions and enhanced work efficiency (Bang et al., 2022; Zhang et al., 2022). Furthermore, governance mechanisms, such as board independence and strong internal controls, mitigate agency costs and streamline coordination, thereby refining the efficiency of decision-making (Katti and Raithatha, 2018).

These strategic and innovative processes, linked to ESG practices, manifest themselves in complex organisational capabilities that competitors find difficult to replicate. A positive and direct impact of ESG factors on market capitalisation has been linked to sound sustainable practices fostered by operational efficiency, reputation enhancement, investor attraction and risk management (Nirino et al., 2021; Bofinger et al., 2022; Janicka and Sajnog, 2022; Zhou et al., 2022).

Thus, innovation, combined with sustainability, constitutes a strategic capability that consolidates efficiency, optimises resource utilisation, strengthens human capital, and refines governance frameworks. As research conducted in recent years has shown, integrating corporate social responsibility (CSR) into core strategies strengthens market positioning (Porter and Kramer, 2006).

The results of recent studies (Lee et al., 2026) show that social investments increase operational efficiency, while environmental and governance investments mainly strengthen market performance. Additionally, corporate strategy moderates these relationships; strategic misalignment, such as rigid governance structures within companies, diminishes ESG effectiveness. However, the same study highlights that European companies are leaders in governance and operational efficiency, suggesting that such companies can aspire to consolidate their market performance.

Moreover, the study by Inaba (2026) reveals that companies listed on representative stock indices in 18 countries around the world, with better environmental, social and governance management levels and higher investment in research and development, were both highly valued on the market and had excellent corporate cash flows in the period 2008-2019.

Other recent studies also demonstrate that companies with superior ESG performance enjoy stronger demand and gain legitimacy from stakeholders, such as investors, which improves operational stability and market performance (Zhou et al., 2022; Chen et al., 2023). Hence, stakeholder theory (Freeman, 1984; Freeman et al., 2017) and technology acceptance models suggest that companies should innovate by adopting sustainable practices to increase user trust and acceptance, which are essential for the successful implementation of sustainable technology and for the appreciation of companies in the market.

Analysis of emerging economies (BRICS) also corroborates the thesis that the impact of ESG performance on market capitalisation is positive and significant, regardless of the specific characteristics of the sector and country (Deng et al., 2025).

In relation to the considerations made, the purpose of this article is to answer the following research question: How does the market evaluate innovative and sustainable companies, including in terms of market capitalisation? (RQ1).

3 Research Methods

The research method adopted in this paper was qualitative (Parker and Northcott, 2016), making use of the multiple case study analysis (Stake, 2005).

In particular, we analysed company and market data, financial statements and sustainability reports for the financial years 2022, 2023 and 2024 of a sample of Italian listed companies selected for their commitment to innovative and sustainable investments, and investigated the relationship between these investments and their market performance.

For this research, we utilised the AIDA database, studying and analysing Italian public limited companies listed on the Milan Stock Exchange. Information regarding market capitalisation was also obtained by consulting studies and analyses conducted by "Borsa Italiana".

The sample comprises the 10 listed Italian companies that have distinguished themselves most for their commitment to supporting cutting-edge, innovative and sustainable investments planned over the period under review, 2022–2024.

These are innovative and sustainable companies that have invested tens or, in most cases, hundreds of millions or billions of euros in the sectors we have investigated, achieving excellent market performance in terms of market capitalisation. They are:

- Enel S.p.A. (energy sector);
- Intesa Sanpaolo S.p.A. (banking sector);
- Leonardo S.p.A. (defence, aerospace and security sector);
- Poste Italiane S.p.A. (postal, logistics, financial, insurance and telecommunications sectors);
- A2A S.p.A (energy, environmental and waste management sectors);
- Ferrari N.V. (automotive sector);
- Webuild S.p.A (construction and engineering sector);
- BF S.p.A (agro-industrial sector);
- Prysmian S.p.A. (manufacturing sector – production of cables for the energy and telecommunications sectors; production of optical fibres);
- Terna S.p.A (energy sector – operator and manager of electricity transmission networks).

4 Results and Discussion

The companies in the sample are highly reputable firms, leaders in their respective markets, including at an international level.

Knowledge-based capabilities enable these companies to effectively communicate their commitment to innovation and ESG: indeed, these investments are included in their 2022–2024 Strategic Plan. This approach raises awareness among investors and the market of the company's potential to achieve greater operational efficiency.

The following table provides an overview of the key findings relating to the investments made by the companies in the sample and the market performance achieved over the three-year period 2022–2024.

Table 1: Key findings

Company	Innov. & Sust. Invest. FY 2022 (€ millions)	Innov. & Sust. Invest. FY 2023 (€ millions)	Innov. & Sust. Invest. FY 2024 (€ millions)	Market Cap. FY 2022 (€ billions)	Market Cap. FY 2023 (€ billions)	Market Cap. FY 2024 (€ billions)
Enel S.p.A.	14,334	14,334	14,334	51.104	68.361	69.925
Intesa Sanpaolo S.p.A.	49.750	49.750	49.750	37.978	48.222	68.693
Leonardo S.p.A.	2,000	750.00	750.00	4.637	8.592	14.918
Poste Italiane S.p.A.	656.750	656.750	656.750	11.851	13.311	17.633

Company	Innov. & Sust. Invest. FY 2022 (€ millions)	Innov. & Sust. Invest. FY 2023 (€ millions)	Innov. & Sust. Invest. FY 2024 (€ millions)	Market Cap. FY 2022 (€ billions)	Market Cap. FY 2023 (€ billions)	Market Cap. FY 2024 (€ billions)
A2A S.p.A.	1,833	1,833	1,833	3.900	5.824	6.720
Ferrari N.V.	880.00	880.00	880.00	36.427	55.064	73.838
Webuild S.p.A.	43.667	52.00	250.00	1.362	1.824	2.839
BF S.p.A.	5.900	8.600	3.200	720.2 millions	1.034	1.163
Prysmian S.p.A.	54.00 (12.60 in R&D)	54.00 (21.50 in R&D)	54.00 (32.80 in R&D)	9.133	11.227	17.823
Terna S.p.A.	2,500	2,000	2,400	13.839	15.152	15.285

Source: Authors

The companies in the sample stand out for their significant commitment to innovative and sustainable investment and for the steady growth in their market capitalisation over the three financial years in question.

Based on the results summarised in the table above, we can make the following observations:

- Enel S.p.A. is the company in the sample that has invested the most in innovation and sustainability over the three financial years in question, totalling approximately €43 billion. The market has welcomed this approach and, indeed, its market capitalisation – with the exception of 2024 (when it was slightly lower than that of Ferrari) – has also been the highest in the sample. The expansion achieved in the 2023 financial year was remarkable, with market capitalisation growing by over 30%, from €51 billion to over €68 billion;
- Intesa Sanpaolo S.p.A. has invested approximately €150 million in innovation and sustainability over the three-year period 2022–2024, thereby achieving excellent market performance. In fact, market capitalisation grew from around 38 billion in 2022 to 48 billion in 2023 and up to 68.7 billion in 2024, recording an increase of over 40% between 2023 and 2024. It is one of the companies with the highest market capitalisation in the sample;
- Leonardo S.p.A, thanks to its investments of €3.5 billion between 2022 and 2024, has been rewarded by the market with an 85% increase in market capitalisation (the highest annual percentage change in the sample) between 2022 and 2023, and growth of over 70% between 2023 and 2024. Consequently, market capitalisation has seen extraordinary growth, rising from €4.6 billion in 2022 to approximately €15 billion in 2024: capitalisation has more than tripled over the course of just two financial years;
- Poste Italiane S.p.A. has invested around 2 billion euros over the three-year period, and its market capitalisation has risen steadily between 2022 and 2024: from 11.8 to 17.6 billion, recording an increase of more than 10% between 2022 and 2023 and an increase of over 30% between 2023 and 2024;
- A2A S.p.A has allocated significant investments to innovation and sustainability, totalling 5.5 billion: it ranked third in the sample in terms of investment volume. Between the 2022 and 2023 financial years, the company recorded an increase of around 50% in its market capitalisation, reaching a market capitalisation of nearly €7 billion in 2024;
- Ferrari N.V. has invested over €2.6 billion in innovation and sustainability, and the market has recognised this commitment: in fact, the company achieved the highest market capitalisation in the sample in the 2024 financial year, amounting to 73.8 billion, whilst its market performance in the 2022 (36.4 billion) and 2023 (55 billion) financial years was the best after that of Enel;
- Webuild S.p.A has allocated over €345 million to innovation and sustainability over the three-year period in question, with progressive investments that have boosted 2023 capitalisation by more than 30% compared with 2022. Following the significant increase in investment in 2024, which was five times that of 2023, capitalisation growth over the same period was close to 60%;
- BF S.p.A increased its investment in innovation and sustainability by over 45% between 2022 and 2023 and, during the same period, its market capitalisation exceeded the one-billion-euro mark, up from €720 million in the 2022 financial year (a 40% increase);
- Prysmian S.p.A. has allocated over €160 million to innovation and sustainability over the three-year period, with investment in research and development set to rise between 2022 and 2024, from €13 million to around €33 million. The market has responded positively to this commitment: in fact, market capitalisation grew by over 20% between 2022 and 2023 and by around 60% between 2023 and 2024;

- Terna S.p.A. ranked second in the sample in terms of the volume of investment allocated to innovation and sustainability, at around €7 billion. Its market capitalisation has grown steadily over the three-year period, exceeding €15 billion as early as 2023.

5 Conclusions

The findings of the study enable us to understand how innovative and sustainable companies, which invest significant resources in these areas, are valued and rewarded by the market, thereby addressing the research question.

Indeed, the market views commitment to research, development, innovation and sustainability positively: this is demonstrated by Enel's market capitalisation, which is the highest in the sample (considering the three-year period 2022–2024), even in light of the higher investments made compared to the other companies in the reference sample. These findings are consistent with the literature (Raimo et al., 2020; Aydogmus et al., 2022; Zhou et al., 2022; Ni and Sun, 2023; Espinosa-Mendez et al., 2024).

The companies in the sample stand out for their prudent management of knowledge-based skills, which is reflected in the inclusion of innovative and ESG investments in their Strategic Plans or in the integration of their Sustainability Plan into their Business Plan (as in the case of Terna). Thanks to this ability to communicate their commitment to innovation and sustainability, companies make investors – even the most sceptical ones – aware of their genuine commitment in these areas.

Enel, by investing €43 billion in innovation and sustainability, has attracted the attention of investors focused on ESG issues. The company estimates that ESG share capital, contributed by socially responsible investors (SRI), will amount to 14.9% in 2022, with an ESG free float of 19.4%; in the 2023 financial year, the ESG share capital stands at 17.5% and the ESG free float at 22.9%; in the 2024 financial year, these figures rise to 23% and 30.1% respectively. At Prysmian too, thanks to the innovative and sustainable investments made, "core ESG" or "ESG-focused" investors account for as much as 40–50% of the share capital. This phenomenon is consistent with the studies conducted by Zhou et al. (2022) and Chen et al. (2023).

Companies that promote investment in innovation and sustainability are viewed by the market as being poised for development and growth, thanks to their greater operational efficiency (as argued by Nirino et al., 2021; Bofinger et al., 2022; Janicka and Sajnog, 2022; Zhou et al., 2022) that they are potentially capable of achieving, as well as thanks to the shared value creation they can deliver. Thus, the direct relationship between investment and operational efficiency, in line also with the research conducted by Lee et al. (2026), fosters positive market expectations that have a "cascading" effect on share prices, causing them to rise.

Market performance is therefore correlated with investment and the integration of ESG measures, as highlighted by Lo and Kwan (2017); Janicka and Sajnog (2022); Zhou et al. (2022); and also by Li et al. (2022) and Inaba (2026).

A steady rise in share prices characterised the performance of all 10 companies in the sample over the three-year period under review. This price growth drove the strongest market performance, as measured by market capitalisation, which increased between 2022 and 2024.

The companies examined, spanning various economic sectors, have all adopted an innovative and sustainable strategy, although their investment choices have focused on different areas.

Some companies, such as Intesa Sanpaolo, have prioritised investments in applied research, artificial intelligence (AI), neuroscience and robotics. Others, such as Leonardo, have aligned their investments with the United Nations Sustainable Development Goals (SDGs); in particular, those relating to decent work and economic growth (SDG 8), industry, innovation and infrastructure (SDG 9) and sustainable cities and communities (SDG 11).

Poste Italiane has focused its investments on automation, modernisation and digital transformation: from electric vehicles to solar panels, from smart post boxes to the technological infrastructure of business support systems.

A2A has decided to focus its investment strategy on the circular economy and energy transition; Ferrari has centred its investment plan on technological innovation, research, environmental sustainability and ecological transition.

Webuild's investments have been directed towards "cleantech", i.e. innovative technologies designed to reduce environmental impact, and subsequently towards research and development.

BF has directed its investments primarily towards research and development, as well as innovative activities aimed at adapting to climate change. Prysmian, too, has channelled most of its investments into research and development and the reduction of GHG emissions.

Terna has focused its investments on reducing CO2 emissions, integrating renewable energy sources, the security and resilience of the electricity system, the digitalisation and technological innovation of the grid (AI, robotics, IoT and BIM), decarbonisation and the energy transition.

These results are very encouraging, as they set an example for other companies (including international ones) that are keen to innovate and operate sustainably, whilst also attracting investors and being recognised and rewarded by the market. Consequently, the market demonstrates its interest in companies that not only invest profitably, but also prove capable of effectively communicating, in their strategic or business plans, the investments they make and the innovative and sustainable objectives they aim to achieve. Investors' positive expectations are based on the potential for growth in operational efficiency, thanks to the investments made, and have an impact on rising share prices and market capitalisation.

The limitations of this research lie in the fact that the analysis is based solely on a study of Italian listed companies, although, to date, the cases examined represent highly significant examples. Future studies aim to identify and analyse other international listed companies that have invested in innovation and sustainability, in order to examine the reaction of foreign markets to such strategic choices.

6 Acknowledgements

Raffaele de Socio: PhD Student, PhD Program in Business Administration, XXXIX Cycle, PNRR DM 118/2023 Generiche, Sapienza University of Rome. This research was carried out as part of the Research Project PNRR 118 Generiche funded by the European Union – NextGenerationEU - scholarship titled: "The effects of the digital and ecological transitions in the annual reports and corporate disclosure".

7 Ethics Declaration

We did not need ethical clearance for the research referred to in this paper.

8 AI Declaration

No artificial intelligence tools were used in this study.

References

Alkaraan, F., Albitar, K., Hussainey, K., Venkatesh, V. (2022). "Corporate transformation toward Industry 4.0 and financial performance: The influence of environmental, social, and governance (ESG)", *Technological Forecasting and Social Change*, Vol. 175, 121423.

Aydogmus, M., Gulay, G., Ergun, K. (2022). "Impact of ESG performance on firm value and profitability", *Borsa Istanbul Review*, Vol. 22, pp. 119-127.

Bang, S. R., Choi, M. C., Ahn, J. Y. (2022). "Human resource practices for corporate social responsibility: Evidence from Korean firms", *Frontiers in Psychology*, Vol.13, 893243.

Barney, J. (1991). "Firm resources and sustained competitive advantage", *Journal of Management*, Vol. 17, No. 1, pp. 99–120.

Bianchi, C. (1998). *Il modello aziendale come modello di economicità*, Kappa, Rome.

Bianchi, C. (2010). *Strutture aziendali nel mercato globalizzato*, Esculapio, Bologna.

Bofinger, Y., Heyden, K. J., Rock, B. (2022). "Corporate social responsibility and market efficiency: Evidence from ESG and misvaluation measures", *Journal of Banking & Finance*, Vol. 134, 106322.

Capaldo, P. (2013). *L'azienda. Centro di produzione*, Giuffrè, Milan.

Chen, S., Song, Y., Gao, P. (2023). "Environmental, social, and governance (ESG) performance and financial outcomes: Analyzing the impact of ESG on financial performance", *Journal of Environmental Management*, Vol. 345, 118829.

Deng, Y., Vitenu-Sackey, P. A., Yang, Z., Liu, X. (2025). "ESG Performance and Market Capitalization of Industries in the BRICS Region: The Mediating Role of Cost of Equity", *Engineering Economics*, Vol. 36, No.5, pp. 580-601.

Espinosa-Méndez, C., Maquieira, C., Arias, J. T. (2024). "ESG performance on the value of family firms: international evidence during Covid-19". *Humanities and Social Sciences Communications*, Vol. 11, No. 1, pp. 1-9.

Fatemi, A., Glaum, M., Kaiser, S. (2018), "ESG performance and firm value: The moderating role of disclosure", *Global finance journal*, Vol. 38, November, pp. 45-64.

Freeman R. E. (1984). *Strategic management: A stakeholder approach*, Cambridge University Press, Cambridge.

Freeman R. E., Kujala J., Sachs S. (2017). *Stakeholder engagement: Clinical research cases*, Springer, Berlin.

García-Amate, A., Ramirez-Orellana, A., Rojo-Ramirez, A. A., Casado-Belmonte, M. P. (2023). "Do ESG controversies moderate the relationship between CSR and corporate financial performance in oil and gas firms?", *Humanities and Social Sciences Communications*, Vol. 10, No. 1, pp. 1-14.

Inaba, K. I. (2026). "Corporate cash value and ESG management: Panel data analyses of stock indices across countries", The North American Journal of Economics and Finance, Vol. 81, 102521.

Janicka, M. and Sajnog, A. (2022). “The ESG Reporting of EU Public Companies. Does the Company's Capitalisation Matter?”, *Sustainability*, Vol. 14, No. 7, p. 4279.

Katti, S. and Raithatha, M. (2018). “Governance practices and agency cost in emerging market: Evidence from India”, *Managerial and Decision Economics*, Vol. 39, No. 6, pp. 712–732.

Lee, F. P., Ting, I. W. K., Lu, W. M. (2026). “Driving Sustainability: ESG and Business Strategy in Global Autonomous Driving Industry”, *Technology in Society*, Vol. 85, 103225.

Li, Z., Feng, L., Pan, Z., Sohail, H. M. (2022). “ESG performance and stock prices: evidence from the COVID-19 outbreak in China”, *Humanities and Social Sciences Communications*, Vol. 9, No. 1, pp. 1-10.

Lo, K. Y. and Kwan, C. L. (2017). “The effect of environmental, social, governance and sustainability initiatives on stock value-Examining market response to initiatives undertaken by listed companies”, *Corporate Social Responsibility and Environmental Management*, Vol. 24, No. 6, pp. 606-619.

Ni, Y. and Sun, Y. (2023). “Environmental, social, and governance premium in Chinese stock markets”, *Global Finance Journal*, Vol. 55, 100811.

Nirino, N., Santoro, G., Miglietta, N., Quaglia, R. (2021). “Corporate controversies and company's financial performance: Exploring the moderating role of ESG practices”, *Technological Forecasting and Social Change*, Vol. 162, 120341.

Parker, L.D. and Northcott, D. (2016). "Qualitative generalising in accounting research: concepts and strategies", *Accounting, Auditing & Accountability Journal*, Vol. 29, No. 6, pp. 1100-1131.

Porter, M. E. and Kramer, M. R. (2006). *The link between competitive advantage and corporate social responsibility*, Harvard Business Review, Vol. 84, No. 12, pp. 78–92.

Raimo, N., de Nuccio, E., Giakoumelou, A., Petruzzella, F., Vitolla, F. (2020). “Nonfinancial information and cost of equity capital: An empirical analysis in the food and beverage industry”, *British Food Journal*, Vol. 123, No. 1, pp. 49-65.

Samy ElDeeb, M., Salah ElGabry, N., Mounir, N., Ahmed, M. (2025). “Nexus among herding behavior, ESG disclosure, and market capitalization in the Egyptian stock market”. *Future Business Journal*, Vol. 11, No. 1, p. 163.

Stake, R. E (2005). *Multiple case study analysis*, Guilford Press, New York.

Vagnoni, E., Cavicchi, C., Borghesi, M., Trincanato, E., Bonnini, S. (2025). “Greener Pathways, strategy, and economic performance: A cross-national study on European service SMEs”, *Business Strategy and the Environment*, Vol. 34, No. 3, pp. 3194–3208.

Wernerfelt, B. (1984). “A resource-based view of the firm”, *Strategic Management Journal*, Vol. 5, No. 2, pp. 171–180.

Wong, C. W. Y., Wong, C. Y., Boon-itt, S. (2020). “Environmental management systems, practices and outcomes: Differences in resource allocation between small and large firms”, *International Journal of Production Economics*, 228.

Zhang, H., Sun, L., & Zhang, Q. (2022). “How workplace social capital affects turnover intention: The mediating role of job satisfaction and burnout”, *International Journal of Environmental Research and Public Health*, Vol. 19, No. 15.

Zhou, G., Liu, L., Luo, S. (2022). “Sustainable development, ESG performance and company market value: Mediating effect of financial performance”, *Business Strategy and the Environment*, Vol. 31, No. 7, pp. 3371-3387.

FedHR5.0: A Human-Centric Federated Learning Framework for Organisational Learning and Workforce Competency Development in Healthcare

Fabio Liberti
University of the Chambers of Commerce, Universitas Mercatorum, Rome, Italy

Abstract

Healthcare organisations face simultaneous pressure to digitalise their workforce, redesign learning processes and adopt artificial intelligence in support of human resource decisions, while protecting sensitive employee data and respecting professional dignity - pressures more acute in healthcare than in other knowledge-intensive sectors, and bearing on clinical and administrative cadres alike.

Three research streams have approached this problem in isolation: federated learning in healthcare, privacy-preserving HR analytics, and organisational learning theory engaging with AI as a partner in knowledge management. Federated learning has predominantly addressed clinical data rather than workforce competency data; HR analytics has not been developed at the intersection with federated infrastructures; organisational learning theory has rarely engaged with the constraints that federated, privacy-preserving learning imposes. We position our contribution explicitly at the three-way intersection.

We introduce FedHR5.0, a conceptual framework for privacy-preserving organisational learning and workforce competency development in healthcare. FedHR5.0 combines federated learning, (ε, δ)-differential privacy and transparent governance with a five-module architecture - trust-based, adaptive learning, emotionally-intelligent assessment, ethical AI, and collaborative knowledge exchange. The framework extends the SECI dynamics of Nonaka and Takeuchi toward a generative-AI-aware GRAI cycle, in which tacit–explicit transitions are mediated by an explicit privacy budget that doubles as a trust-building instrument. The artefact is positioned as a Design Science Research *improvement*-type contribution.

Empirical validation is scoped as future work and given a concrete pathway: a multi-hospital pilot, four operationally measurable evaluation axes (model accuracy, effective privacy budget, organisational readiness, worker participation) and a three-phase iterative protocol grounded in Design Science Research.

For HR practice, the framework repositions learning and development as a strategic driver of human–AI collaboration and treats workforce participation as a structural design requirement, applicable to clinical and administrative cadres alike. For research, it offers a starting point for cumulative design knowledge at the emerging intersection of federated learning, organisational learning and human-centric HR analytics.

Keywords: Organisational learning; Federated learning; Healthcare workforce; AI-based learning systems; Workforce transformation.

Paper type: Academic Research Paper.

1 Introduction

Healthcare organisations face simultaneous pressure to digitalise their workforce, redesign learning processes and adopt artificial intelligence in support of human resource decisions, while protecting sensitive employee data and respecting professional dignity. Demographic ageing, accelerated digital health transformation and the diffusion of AI across people management create a constellation of constraints more acute in healthcare than in any other knowledge-intensive sector (Pavuluri et al. 2024, de Kok et al. 2023).

Three research streams have approached this problem in isolation: federated learning in healthcare has matured rapidly but addresses primarily clinical data rather than workforce competency data (Rieke et al. 2020, Chaddad et al. 2024); privacy-preserving HR analytics has begun to engage with transparency and accountability (Tursun-bayeva et al. 2018, Yanamala 2024), yet not at the intersection with federated infrastructures; organisational learning theory has engaged with AI in knowledge management (Jarrahi et al. 2023, Böhm & Durst 2025), but rarely with the architectural constraints federated learning imposes.

We position our contribution within healthcare workforces in their full breadth - clinical and administrative - recognising that recent evidence on health service management workforce capacity (Brommeyer et al. 2022) and on AI's reshaping of health information management (Stanfill & Marc 2019) positions administrative cadres as primary subjects, not residual recipients. Clinical staff exemplify the most acute case of tacit-knowledge dimensions; administrative roles extend the same logic to digital literacy, data governance and change-management capacity.

This paper asks how an integrated, human-centric architecture can support privacy-preserving organisational learning and workforce competency development across networks of healthcare institutions. Our answer is FedHR5.0: a conceptual framework combining federated learning (McMahan et al. 2016), (ε, δ)-differential privacy and transparent governance with a five-module architecture - trust-based, adaptive learning, emotionally-intelligent assessment, ethical AI, and collaborative knowledge exchange. The framework extends the SECI dynamics of Nonaka & Takeuchi (2009) toward a GRAI cycle (Böhm & Durst 2025), in which tacit–explicit transitions are mediated by an explicit privacy budget that doubles as a trust-building instrument.

Section 2 reviews background; Section 3 positions the contribution; Section 4 articulates the framework; Section 5 details methodology; Section 6 discusses implications; Section 7 outlines a validation pathway; Section 8 concludes.

2 Background

2.1 Organisational learning and workforce transformation in healthcare

Healthcare organisations are knowledge-intensive, high-reliability environments in which learning capability is a precondition for quality and safety. The Learning Health System frame consolidates around governance of the learning cycle, staff engagement, data infrastructure, analytical capacity and action loops, with enabling conditions for organisations to act on what they observe (Easterling et al. 2021). The scoping review by de Kok et al. (2023) identifies five attributes - leadership, culture, team development, change management and customer orientation - that distinguish healthcare organisations with mature learning capabilities from those without.

At the individual and team layer, quality-improvement evidence shows that sustainability of change in healthcare depends on learning processes that reach the mental models of practitioners, in line with the classical reading of double-loop learning (Hovlid et al. 2012).

2.2 Reskilling and upskilling in AI-driven contexts

The diffusion of artificial intelligence across organisational functions has reopened debate on reskilling, upskilling and the strategic role of the learning and development function. Jarrahi et al. (2023) argue that AI value materialises only when algorithmic capabilities are co-developed with human learning processes, through internal routinisation and outward-oriented strategic learning. In this reading, the learning function is no longer a support activity but a precondition for AI-driven transformation.

At the operational layer, systematic mappings document AI-enabled adaptive learning systems with personalised content sequencing, real-time feedback and performance analytics (Kabudi et al. 2021, Gligorea et al. 2023).

The transfer of these insights to healthcare workforce development is not automatic. Healthcare workforces - clinical and administrative - combine codified procedural knowledge with tacit, relational and emotional competencies acquired through practice and mentorship. Clinical staff exemplify this duality most acutely, but it extends to administrative and management roles, where digital literacy, data governance and change-management capacity are increasingly tacit and contextual. Recent scoping evidence on health service management workforce capacity confirms that digital health and informatics are reshaping the competencies expected of administrative cadres (Brommeyer et al. 2022), while AI is transforming the role of health information management professionals at the boundary between clinical and administrative work (Stanfill & Marc 2019).

2.3 Privacy, trust and human-centric AI in HR

The integration of AI into HR processes brings its own tensions. People analytics has matured as a research field with explicit value propositions across the employment cycle (Tursunbayeva et al. 2018); its dark sides - surveillance concerns, erosion of employee trust - have been systematised with equal attention (Giermindl et al. 2021). Practitioner-oriented work highlights transparency, privacy and accountability as structural conditions for legitimacy (Yanamala 2024).

A parallel debate frames AI and human judgement as augmentation rather than automation: Raisch & Krakowski (2021) articulate the automation–augmentation paradox, while Fügener et al. (2021) examine human-in-the-loop arrangements where the human risks becoming a downstream executor. AI design for HR must therefore protect human judgement, the relational texture of workplace learning, and employee agency over their data.

These three layers - organisational learning capability, AI-driven reskilling, and the tension between trust and algorithmic HR - justify the three-way intersection occupied by our framework.

3 Related Work

The contribution of this paper sits at the intersection of three streams that have so far remained largely disjoint: federated learning in healthcare, privacy-preserving analytics with a focus on human-resource data, and organisational learning theory engaging with artificial intelligence in competency development. The subsections below review each stream and prepare the positioning of the framework.

3.1 Federated learning in healthcare

Federated learning emerged as a distributed training paradigm in which participants collaboratively build a shared model while keeping data local (McMahan et al. 2016), with subsequent work systematising open challenges around communication cost, heterogeneity and privacy (Li et al. 2019).

In healthcare, cross-institutional model training has been framed as a response to the privacy and regulatory constraints that fragment clinical data (Rieke et al. 2020). Surveys have mapped clinical applications and clarified architectural choices for heterogeneous settings (Chaddad et al. 2024); recent reviews integrate federated learning with privacy-preserving techniques and IoT infrastructures (Abbas et al. 2024); and personalised variants address the non-IID data of multi-hospital networks (Tan et al. 2021).

Two scope limits matter here: federated learning has predominantly addressed clinical rather than workforce data, and contributions specifically targeting competency development in healthcare organisations are scarce.

3.2 Privacy-preserving HR analytics

A parallel stream investigates analytics on human-resource data. The scoping review by Tursunbayeva et al. (2018) clarifies the conceptual boundaries of people analytics, while Giermindl et al. (2021) survey its dark sides - surveillance, erosion of trust - and argue that technical adequacy alone does not warrant legitimacy.

The intersection of federated learning with HR analytics is remarkably under-developed: recent practitioner-oriented work discusses transparency and accountability in AI-enhanced HR processes (Yanamala 2024), and a consolidated governance framework for federated learning has been articulated (Eden et al. 2025). A very recent contribution applies federated learning to a single retention-prediction task on benchmark HR data (B. 2026), but does not propose an integrated architecture and remains task-specific. None of these contributions targets workforce competency data as a framework-level concern.

3.3 Organisational learning, AI-based learning and competency development

The third stream draws on organisational learning theory and on the literature examining how AI reshapes the processes through which organisations create, retain and circulate knowledge. The reference is the knowledge-creating company of Nonaka & Takeuchi (2009), whose four modes of conversion - socialisation, externalisation, combination, internalisation - describe how tacit and explicit knowledge interact within an organisation, and which we read as a substrate liftable to inter-organisational settings when suitable infrastructures exist.

Recent work frames the human–AI relationship in knowledge management as partnership rather than substitution. Jarrahi et al. (2022) map complementary roles along the four classical KM dimensions; Jarrahi et al. (2023) develop the notion of an artificial capital operating in synergy with human capital through reciprocal learning. Böhm & Durst (2025) propose the GRAI extension of SECI, in which generative AI operates as an additional agent within the knowledge-creation spiral, legitimising an inter-organisational reading of knowledge conversion.

At the practical layer, healthcare-specific competency frameworks remain heterogeneous and non-standardised (Batt et al. 2019), while explainable AI components - particularly relevant to trust - have been systematised under the Healthcare 5.0 banner (Saraswat et al. 2022).

The three streams converge on the questions motivating our framework yet remain largely disjoint: federated learning rarely engages with organisational learning theory, organisational learning scholarship seldom discusses federated architectures, and HR analytics has not been developed at the intersection with federated learning for competency development. Positioning FedHR5.0 at this three-way intersection is the genuine novelty we claim.

4 Conceptual Framework: FedHR5.0

4.1 Rationale and overview

FedHR5.0 is a theoretical framework combining federated learning (McMahan et al. 2016, Li et al. 2019), differential privacy and transparent governance to enable privacy-preserving organisational learning in healthcare, where institutions face the twin constraints of protecting sensitive employee data and preserving human dignity in algorithmic HR decisions (Pavuluri et al. 2024, de Kok et al. 2023).

The framework addresses three intertwined challenges: supporting cross-institutional learning on workforce skills without centralising sensitive data; designing AI-based learning systems that enhance rather than replace human judgement (Raisch & Krakowski 2021, Fügener et al. 2021); embedding trust, transparency and employee participation in AI-driven HR practices (Yanamala 2024, Tursunbayeva et al. 2018).

Conceptually, FedHR5.0 reframes federated learning as infrastructure for collaborative HR analytics across health-care organisations (Rieke et al. 2020, Chaddad et al. 2024): institutions train local models on workforce data and share only model updates, while the aggregated global model powers distributed learning platforms for talent management, skill assessment and personalised training.

4.2 A five-module architecture

At the core of FedHR5.0 lies a five-module architecture that operationalises human-centric HR transformation in healthcare. Each module addresses a distinct layer of the socio-technical system - trust, learning delivery, competency assessment, ethical guardrails, and inter-organisational knowledge exchange - and interacts with the others through continuous feedback (Figure 1).

Trust-Based Architecture. The first module establishes the conditions for trust in algorithmic HR decisions. It combines explainable AI components (Saraswat et al. 2022), auditable decision trails and participatory design practices that involve workers in system configuration. Privacy guarantees are communicated to employees not only as technical features but as intelligible commitments over their own data. Governance mechanisms follow the taxonomy of Eden et al. (2025) - procedural, relational and structural elements - to make accountability visible across the federation.

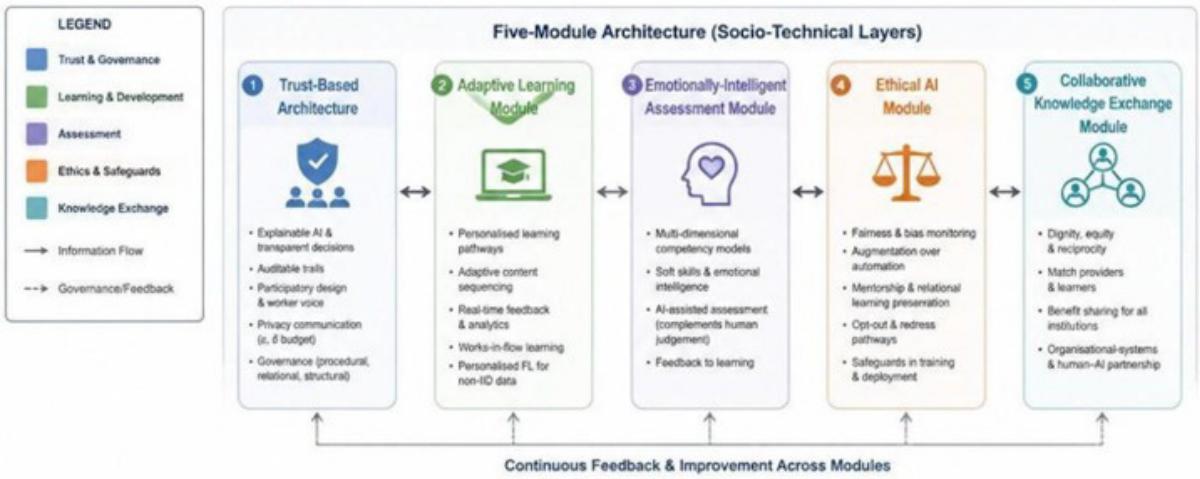


Figure 1: The FedHR5.0 five-module architecture across socio-technical layers, with bidirectional feedback and improvement loops between modules

Adaptive Learning Module. The second module delivers AI-driven personalised learning pathways. It builds on systematic evidence on adaptive learning in AI-enabled educational settings (Kabudi et al. 2021, Glig-orea et al. 2023) and on personalised federated techniques accommodating non-IID data across organisations (Tan et al. 2021). The module provides real-time feedback, adaptive content sequencing and performance analytics embedded in the flow of daily work, supporting continuous learning across heterogeneous clinical environments.

Emotionally-Intelligent Assessment Module. The third module extends competency models beyond cognitive and technical dimensions. It incorporates soft skills, emotional intelligence and relational capabilities into the assessment framework, recognising the centrality of tacit knowledge in healthcare practice (Hoge et al. 2005). The module is informed by recent evidence on AI-assisted evaluation of soft skills (Gafni et al. 2024).

Assessment complements rather than replaces peer and supervisor judgement, and feeds signals back into the adaptive learning module.

Ethical AI Module. The fourth module focuses on ethical guardrails. It embeds fairness constraints into recruitment, performance management and development recommendations; monitors algorithmic bias and deskilling dynamics (Pavuluri et al. 2024, Giermindl et al. 2021); and preserves mentorship and relational learning as core elements of professional development. The design follows the augmentation-over-automation orientation (Raisch & Krakowski 2021): AI extends rather than substitutes human judgement and peer learning. Safeguards are built into both the training pipeline and the deployment layer, including mentorship preservation indicators and opt-out pathways.

Collaborative Knowledge Exchange Module. The fifth module governs the exchange of knowledge across institutions. It prioritises dignity, equity and reciprocity in matching knowledge providers and learners, drawing on organisational-systems approaches to cultural competence (McCalman et al. 2017) and on the human–AI partnership view of knowledge management (Jarrahi et al. 2023). The module also addresses a risk identified in governance reviews of federated learning (Eden et al. 2025): smaller or less-resourced institutions must retain meaningful access to the federated signal, rather than contributing data without benefiting from aggregated insights.

4.3 Extending the SECI Model to Distributed Privacy-Preserving Contexts

From a knowledge management standpoint, FedHR5.0 extends the SECI model (Nonaka & Takeuchi 2009) to a distributed, privacy-preserving setting: the four modes of conversion - originally intra-organisational - are lifted to an inter-organisational space, with local models carrying externalised and combined knowledge sharable without exposing individual-level data. This is aligned with the GRAI logic of Böhm & Durst (2025), in which generative AI operates as an additional agent in the knowledge-creation spiral; differential privacy formalises what leaves each organisation, while federated aggregation re-injects refined knowledge.

5 Methodology

5.1 A design science research approach

We frame the construction of FedHR5.0 within the Design Science Research tradition, following the iterative procedure of Peffers et al. (2007). Within the typology of DSR contributions (Gregor & Hevner 2013), FedHR5.0 is positioned as an *improvement*: the underlying problem is known, but the combined architecture is new. Documentation of design choices follows Brocke et al. (2020) on cumulative design knowledge.

5.2 Formalisation: federated learning with differential privacy

At the technical core, FedHR5.0 relies on standard federated averaging (McMahan et al. 2016). Each healthcare organisation k maintains a local dataset D_k of workforce data; at round t, local models ϑ^t are updated via stochastic gradient descent on D_k and aggregated into a global model k

$$\theta^{t+1} = \sum_k \frac{n_k}{n} \theta_k^t, \qquad n_k = |D_k|, \quad n = \sum_k n_k .$$

The protocol inherits the open challenges systematised in Li et al. (2019), and adopts personalised variants (Tan et al. 2021) to accommodate non-IID distributions across institutions. Figure 2 illustrates the full pipeline, including the differential privacy step described below.

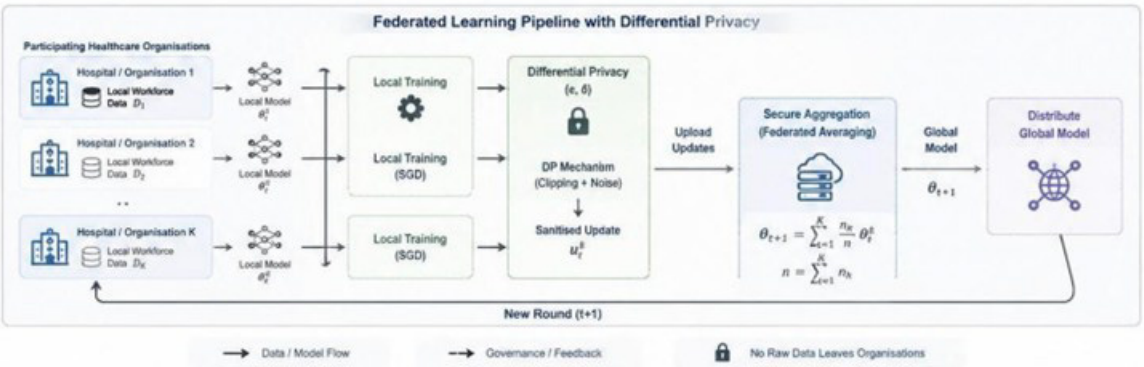


Figure 2: Federated learning pipeline with (ε, δ)-differential privacy: local training on each organisation's workforce data; updates sanitised via clipping and noise; federated averaging; global model broadcast back. No raw data leaves any organisation.

On top of federated aggregation, FedHR5.0 applies (*ε, δ*)-differential privacy to the local model updates before they leave the organisation, in line with the protocols documented in recent reviews of privacy-preserving FL in healthcare (Abbas et al. 2024).

The choice to apply differential privacy *within* each organisation, rather than relying on federated learning alone, deserves explicit justification. Federated learning already prevents raw personal data from crossing organisational boundaries, but local model updates still encode statistical information about individual employees and are potentially vulnerable to inference attacks. Differential privacy provides a formal, quantifiable guarantee on the information any single update can reveal, and makes this guarantee communicable to the workforce as an explicit privacy budget. In FedHR5.0, therefore, differential privacy is not a redundant technical layer, but a *trust-building instrument*: it converts an implicit protection ("your data does not leave the hospital") into an auditable commitment ("the information derived from your data is bounded by *ε, δ*"). This positioning is what allows the trust-based module of the framework (Section 4.2.1) to engage the worker as a participating agent rather than a residual beneficiary.

5.3 Rigour and iteration

We pursue rigour along three axes - utility, feasibility, novelty - per Gregor & Hevner (2013). Utility and feasibility are addressed conceptually here; novelty is argued via the three-way intersection of Section 3. Empirical validation is scoped to Section 7 as a structured pathway.

6 Discussion and Implications

6.1 Theoretical positioning and analytical stance

Our analysis suggests that FedHR5.0 contributes at the three-way intersection of federated learning in healthcare, privacy-preserving HR analytics and human-centric organisational learning, qualifying as an *improvement* contribution (Gregor & Hevner 2013). The framework extends the SECI dynamics of Nonaka & Takeuchi (2009) toward a GRAI-style cycle in which tacit–explicit transitions are mediated by an explicit privacy budget (Böhm & Durst 2025).

6.2 Implications for workforce transformation

The framework bears tangible implications for workforce transformation. First, competency development can be redesigned around multi-dimensional signals: the emotionally-intelligent assessment module (Section 4.2.3) invites HR functions to complement output-based metrics with soft skills, tacit knowledge and emotional intelligence, countering one-dimensional risks of algorithmic HR (Giermindl et al. 2021). Second, personalisation of learning pathways via the adaptive learning module can reasonably be expected to strengthen retention by tying everyday learning practices to employee empowerment and engagement. Third, the framework repositions HR from support activity to strategic driver of human–AI collaboration (Jarrahi et al. 2023), with line managers and CHROs as primary design stakeholders rather than downstream recipients. Fourth, relational learning practices - mentorship, peer learning, communities of practice - enter the architecture as a structural design requirement (Raisch & Krakowski 2021).

6.3 Implementation conditions, readiness, and trade-offs

The framework's actionability rests on a set of organisational preconditions. Drawing on the five attributes of mature learning capabilities (de Kok et al. 2023) - leadership, culture, team development, change management and customer orientation - a federated deployment can be characterised by an implementation checklist that translates each attribute into a concrete precondition, from visible leadership sponsorship to demonstrable change-management capacity. Organisational readiness can be observed along four measurable factors: maturity of data governance, absence of internal information silos, availability of edge or cloud computational resources, and alignment with GDPR and participatory design practices (Pavuluri et al. 2024). Alongside readiness, an explicit register of risks and trade-offs keeps design choices auditable: technically, the tension between privacy budget and global model utility, and between personalisation and equity; organisationally, inter-institutional coordination costs, residual information leakage through model updates, and the risk of eroding worker dignity when transparency is nominal (Fügener et al. 2021). Sustainability of change is itself an observable variable (Hovlid et al. 2012); governance arrangements grounded in the taxonomy of Eden et al. (2025) provide the instrumentation to keep these trade-offs visible.

6.4 Methodological implications and research agenda

These choices raise open methodological questions for DSR in HR contexts: evaluation criteria that reconcile technical privacy budgets with organisational trust metrics, and longitudinal designs to verify augmentation claims over time. Section 7 outlines a concrete validation pathway.

7 Limitations and Future Work

This paper presents a conceptual framework; empirical validation is the principal limitation of the present contribution. Rather than deferring evaluation to an open horizon, we outline a concrete pathway in line with the reviewers' guidance.

Setting. We propose a multi-hospital network in Italy as the primary empirical setting, with an initial cohort of three to five medium-sized institutions (500–1,500 employees each) engaged in workforce competency development. Geographic and dimensional heterogeneity across the network is an asset rather than a confound: it stresses the non-IID robustness that the framework is designed to address.

Evaluation criteria. We propose four operationally measurable axes: (i) global-model accuracy on the workforce-analytics tasks the framework enables, evaluated through standard multi-task machine-learning metrics; (ii) effective privacy budget, verified empirically against the declared (ε, δ) targets under realistic training loads; (iii) an organisational readiness score grounded in the attributes identified by de Kok et al. (2023), measured ex-ante and ex-post via a structured instrument; (iv) a worker participation index combining opt-in rate, consultation of the privacy budget, and engagement with participatory design moments (Easterling et al. 2021).

Methodological approach. A three-phase pilot with iterative DSR cycles (Peffers et al. 2007): (i) readiness assessment and pre-deployment (preconditions verification, federated infrastructure setup, workforce sensitisation); (ii) deployment and observation over six to twelve months on a sub-network, with parallel technical and organisational data collection; (iii) evaluation and design refinement for the next iteration.

8 Conclusions

FedHR5.0 is a conceptual framework for privacy-preserving organisational learning and workforce competency development in healthcare, designed to operate at the three-way intersection of federated learning, human-centric HR analytics and organisational learning theory. Its distinctive contribution lies in the integration of a five-module architecture - trust-based, adaptive learning, emotionally-intelligent assessment, ethical AI, and collaborative knowledge exchange - over a federated-differential-privacy substrate, oriented to clinical and administrative cadres alike, and rooted in an updated reading of the SECI model for the age of generative artificial intelligence (Nonaka & Takeuchi 2009, Böhm & Durst 2025).

Methodologically, the artefact is positioned as an *improvement*-type contribution within the Design Science Research tradition (Gregor & Hevner 2013, Peffers et al. 2007). Empirical validation is scoped as future work and given a concrete pathway in Section 7, with a multi-hospital pilot, four operationally measurable evaluation axes, and a three-phase iterative protocol.

Beyond the immediate artefact, FedHR5.0 invites the research community to accumulate design knowledge at this emerging intersection, in line with the spirit of cumulative design research (Brocke et al. 2020).

9 Ethics declaration

This paper presents a conceptual framework and does not report original empirical research involving human participants, identifiable personal data, or animal subjects. Consequently, no ethical clearance was required for the research described herein.

10 AI declaration

AI-based tools were used selectively during the preparation of this manuscript to support the organisation of the bibliographic synthesis and the preliminary drafting of sections of the text. Any generated content has been verified for accuracy and relevance, edited, and revised by the author, who retains full accountability for the scientific and editorial decisions reflected in the final manuscript.

References

Abbas, S. R., Abbas, Z., Zahir, A. & Lee, S. (2024), 'Federated learning in smart healthcare: A comprehensive review on privacy, security, and predictive analytics with iot integration', *Healthcare* . DOI: 10.3390/healthcare12242587.

B., B. (2026), 'Enhancing employee retention prediction through federated learning and associative classification using fedcba: A case study on ibm hr analytics', *Arabian Journal for Science and Engineering* . DOI: 10.1007/s13369-026-11252-8.

Batt, A., Tavares, W. & Williams, B. (2019), 'The development of competency frameworks in healthcare professions: a scoping review', *Advances in Health Sciences Education* . DOI: 10.1101/19003475.

Brocke, J., Winter, R., Hevner, A. & Maedche, A. (2020), 'Special issue editorial - accumulation and evolution of design knowledge in design science research: A journey through time and space', *J. Assoc. Inf. Syst.* . DOI: 10.17705/1jais.00611.

Brommeyer, M., Whittaker, M., Mackay, M., Ng, F. & Liang, Z. (2022), 'Building health service management workforce capacity in the era of health informatics and digital health - a scoping review', *International journal of medical informatics* . DOI: 10.1016/j.ijmedinf.2022.104909.

Böhm, K. & Durst, S. (2025), 'Knowledge management in the age of generative artificial intelligence – from seci to grai', *VINE Journal of Information and Knowledge Management Systems* . DOI: 10.1108/vjikms-10-2024-0357. Chaddad, A., Wu, Y. & Desrosiers, C. (2024), 'Federated learning for healthcare applications', *IEEE Internet of Things Journal* . DOI: 10.1109/jiot.2023.3325822.

de Kok, K., van der Scheer, W. K., Ketelaars, C. & Leistikow, I. (2023), 'Organizational attributes that contribute to the learning & improvement capabilities of healthcare organizations: a scoping review', *BMC Health Services Research* . DOI: 10.1186/s12913-023-09562-w.

Easterling, D., Perry, A. C., Woodside, R., Patel, T. & Gesell, S. (2021), 'Clarifying the concept of a learning health system for healthcare delivery organizations: Implications from a qualitative analysis of the scientific literature', *Learning Health Systems* . DOI: 10.1002/lrh2.10287.

Eden, R., Chukwudi, I., Bain, C., Barbieri, S., Callaway, L., de Jersey, S. J., George, Y., Gorse, A.-D., Lawley, M., Marendy, P., McPhail, S. M., Nguyen, A., Samadbeik, M. & Sullivan, C. M. (2025), 'A scoping review of the governance of federated learning in healthcare', *NPJ Digital Medicine* . DOI: 10.1038/s41746-025-01836-3.

Fügener, A., Grahl, J., Gupta, A. & Ketter, W. (2021), 'Will humans-in-the-loop become borgs? merits and pitfalls of working with ai', *MIS Q.* . DOI: 10.25300/misq/2021/16553.

Gafni, R., Aviv, I., Kantsepolsky, B., Sherman, S., Rika, H., Itzkovich, Y. & Barger, A. (2024), 'Objectivity by design: The impact of ai-driven approach on employees' soft skills evaluation', *Inf. Softw. Technol.* . DOI: 10.1016/j.infsof.2024.107430.

Giermindl, L., Strich, F., Christ, O., Leicht-Deobald, U. & Redzepi, A. (2021), 'The dark sides of people analytics: reviewing the perils for organisations and employees', *European Journal of Information Systems* . DOI: 10.1080/0960085x.2021.1927213.

Gligorea, I., Cioca, M., Oancea, R., Gorski, A., Gorski, H. & Tudorache, P. (2023), 'Adaptive learning using artificial intelligence in elearning: A literature review', *Education Sciences* . DOI: 10.3390/educsci13121216.

Gregor, S. & Hevner, A. (2013), 'Positioning and presenting design science research for maximum impact', *MIS Q.* . DOI: 10.25300/misq/2013/37.2.01.

Hoge, M., Paris, M., Adger, H., Collins, F., Finn, C. V., Fricks, L., Gill, K., Haber, J., Hansen, M., Ida, D., Kaplan, L., Northey, W. F., O'Connell, M., Rosen, A. L., Taintor, Z., Tondora, J. & Young, A. (2005), 'Workforce competencies in behavioral health: An overview', *Administration and Policy in Mental Health and Mental Health Services Research* . DOI: 10.1007/s10488-005-3259-x.

Hovlid, E., Bukve, O., Haug, K., Aslaksen, A. & von Plessen, C. (2012), 'Sustainability of healthcare improvement: what can we learn from learning theory?', *BMC Health Services Research* . DOI: 10.1186/1472-6963-12-235.

Jarrahi, M. H., Askay, D. A., Eshraghi, A. & Smith, P. (2022), 'Artificial intelligence and knowledge management: A partnership between human and ai', *Business Horizons* . DOI: 10.1016/j.bushor.2022.03.002.

Jarrahi, M. H., Kenyon, S., Brown, A., Donahue, C. & Wicher, C. (2023), 'Artificial intelligence: a strategy to harness its power through organizational learning', *Journal of Business Strategy* . DOI: 10.1108/jbs-11-2021-0182.

Kabudi, T., Pappas, I. & Olsen, D. H. (2021), 'Ai-enabled adaptive learning systems: A systematic mapping of the literature', *Comput. Educ. Artif. Intell.* . DOI: 10.1016/j.caeai.2021.100017.

Li, T., Sahu, A. K., Talwalkar, A. & Smith, V. (2019), 'Federated learning: Challenges, methods, and future directions', *IEEE Signal Processing Magazine* . DOI: 10.1109/msp.2020.2975749.

McCalman, J., Jongen, C. & Bainbridge, R. (2017), 'Organisational systems' approaches to improving cultural competence in healthcare: a systematic scoping review of the literature', *International Journal for Equity in Health* . DOI: 10.1186/s12939-017-0571-5.

McMahan, H. B., Moore, E., Ramage, D., Hampson, S. & y Arcas, B. A. (2016), 'Communication-efficient learning of deep networks from decentralized data', Proceedings of the 20 th International Conference on Artificial Intelligence and Statistics (AISTATS) 2017. JMLR: W&CP volume 54.

Nonaka, I. & Takeuchi, H. (2009), 'The knowledge-creating company', *Economic Impact of Knowledge* . DOI: 10.1016/0024-6301(96)81509-3.

Pavuluri, S. K., Sangal, R. B., Sather, J. & Taylor, A. (2024), 'Balancing act: the complex role of artificial intelligence in addressing burnout and healthcare workforce dynamics', *BMJ Health & Care Informatics* . DOI: 10.1136/bmjhci-2024-101120.

Peffers, K., Tuunanen, T., Rothenberger, M. & Chatterjee, S. (2007), 'A design science research methodology for information systems research', *Journal of Management Information Systems* . DOI: 10.2753/mis0742-1222240302.

Raisch, S. & Krakowski, S. (2021), 'Artificial intelligence and management: The automation–augmentation paradox', *Academy of Management Review* . DOI: 10.5465/amr.2018.0072.

Rieke, N., Hancox, J., Li, W., Milletarì, F., Roth, H., Albarqouni, S., Bakas, S., Galtier, M., Landman, B., Maier-Hein, K. H., Ourselin, S., Sheller, M. J., Summers, R. M., Trask, A., Xu, D., Baust, M. & Cardoso, M. (2020), 'The future of digital health with federated learning', *NPJ Digital Medicine* . DOI: 10.1038/s41746-020-00323-1.

Saraswat, D., Bhattacharya, P., Verma, A., Prasad, V., Tanwar, S., Sharma, G., Bokoro, P. & Sharma, R. (2022), 'Explainable ai for healthcare 5.0: Opportunities and challenges', *IEEE Access* . DOI: 10.1109/access.2022.3197671.

Stanfill, M. & Marc, D. T. (2019), 'Health information management: Implications of artificial intelligence on healthcare data and information management', *Yearbook of Medical Informatics* . DOI: 10.1055/s-0039-1677913.

Tan, A., Yu, H., zhen Cui, L. & Yang, Q. (2021), 'Towards personalized federated learning', *IEEE Transactions on Neural Networks and Learning Systems* . DOI: 10.1109/tnnls.2022.3160699.

Tursunbayeva, A., Lauro, S. & Pagliari, C. (2018), 'People analytics - a scoping review of conceptual boundaries and value propositions', *Int. J. Inf. Manag.* . DOI: 10.1016/j.ijinfomgt.2018.08.002.

Yanamala, K. K. R. (2024), 'Transparency, privacy, and accountability in ai-enhanced hr processes', *Journal of Advanced Computing Systems* . DOI: 10.69987/jacs.2023.30302.

INSPIRE PM as a Human–AI Hybrid Knowledge Ecosystem for Sustainable Project-Based Organizations

Antonio Bassi
Association Project Management Switzerland

Abstract

Project-based organizations are increasingly exposed to a combination of uncertainty, accelerated technological change, and rising sustainability expectations. In this context, many organizations are investing in digital platforms, analytics capabilities, and artificial intelligence applications, yet they often struggle to transform dispersed project information into reliable and actionable knowledge. The spread of generative AI has intensified this tension: while organizations can now generate answers, summaries, and recommendations at scale, they frequently lack adequate governance mechanisms for knowledge provenance, accountability, privacy, and ethical use. Starting from the INSPIRE PM framework and its broader grounding in Project Management Evolutive (PME), this paper argues that sustainable project performance depends on the ability to build Human–AI Hybrid Knowledge Ecosystems (HA-HKE), that is, socio-technical configurations in which human actors and AI systems collaboratively generate, validate, circulate, and apply project knowledge under explicit governance rules. The paper develops a conceptual architecture for such ecosystems, identifies the key knowledge assets and flows required across the project lifecycle, and proposes a design-oriented research agenda centered on three artifacts: a hybrid knowledge ecosystem blueprint, a decision intelligence loop, and a lightweight governance checklist for responsible GenAI use in project contexts. The paper contributes conceptually by reframing project management as a knowledge ecosystem challenge, methodologically by outlining a design science approach with multiple-case evaluation, and managerially by offering a practical pathway for PMOs and project leaders seeking to combine AI-enabled decision support with sustainability, traceability, and human accountability.

Keywords: Human–AI hybrid knowledge; Knowledge ecosystems; Sustainable project management; Decision Intelligence; Project Management Evolutive (PME)

Paper Type: Academic Research Paper

1 Introduction

Project-based organizations are operating in an environment that is both richer in data and poorer in coherence. On the one hand, projects now generate an unprecedented volume of digital traces: schedules, reports, risk registers, chats, forecasts, dashboards, lessons learned, sustainability metrics, decisions, approvals, exceptions, and stakeholder feedback. On the other hand, the ability to turn this fragmented informational landscape into trustworthy knowledge remains limited. Many organizations still rely on disconnected repositories, isolated tools, weak knowledge governance, and ad hoc interpretive practices. As a consequence, project teams often have access to more information than ever before, but not necessarily to better judgment, better alignment, or better decisions.

This challenge is becoming more acute with the rapid diffusion of generative AI. The problem is no longer access to answers, but the governance of the answers that are produced. In project environments, GenAI can summarize documents, generate reports, suggest actions, classify risks, or support scenario analysis. Yet the expansion of AI-supported output does not automatically imply an increase in knowledge quality. Without adequate controls, AI can accelerate ambiguity, produce plausible but weakly grounded recommendations, obscure provenance, and dilute accountability. What appears as “intelligent support” can easily become a new layer of opacity if organizations do not define who validates what, how outputs are traced, how privacy is protected, and where human responsibility remains non-transferable.

The starting point of this paper is that the central issue is not merely one of technological adoption. Rather, it concerns the design of a socio-technical and organizational environment in which knowledge can be created, validated, and reused in ways that improve project performance while remaining aligned with sustainability objectives. In this regard, the INSPIRE PM framework offers an especially useful lens because it is conceived not as a rigid methodology but as a flexible architecture capable of integrating technology, sustainability, predictive analytics, innovation, responsiveness, and efficiency into project management practice. The broader conceptual grounding in Project Management Evolutive (PME) further reinforces this view by presenting projects as dynamic ecosystems where learning, adaptability, data-informed decision-making, and integrated sustainability are foundational.

Within this perspective, the paper develops the concept of **Human–AI Hybrid Knowledge Ecosystems (HA-HKE)** for project-based organizations. The argument advanced here is that sustainable project performance depends on the organizational capability to orchestrate knowledge assets across a hybrid ecosystem in which human actors and AI systems collaborate under explicit governance rules. This requires moving beyond a "tool adoption" narrative toward a genuine ecosystem design problem: organizations must determine which knowledge matters, where it is generated, how it is validated, how it circulates, who is accountable for it, and how it is linked to sustainability-related decisions.

The paper is guided by three research questions already identified in the extended abstract:

- ***RQ1.*** *What knowledge assets and knowledge flows are essential to enable effective human–AI collaboration across the project lifecycle in a sustainable organization?*
- ***RQ2.*** *Which governance mechanisms are needed to ensure trustworthy AI-augmented project knowledge without undermining human accountability?*
- ***RQ3.*** *How can organizations measure the value of hybrid knowledge ecosystems in terms of project performance, resilience, and ESG outcomes?*

The remainder of the paper addresses these questions by first clarifying the theoretical positioning of the study, then articulating the concept of hybrid knowledge ecosystems, and finally proposing a design-oriented research architecture capable of translating the conceptual model into practical organizational artifacts.

2 From Project Knowledge to Hybrid Knowledge Ecosystems

Project management has always been deeply knowledge-intensive, even when this dimension was not made explicit. Projects are not executed only through tasks and deliverables, but through interpretation, judgment, coordination, negotiation, prioritization, and learning. The formal side of project knowledge is usually visible in plans, schedules, risk logs, requirements, budgets, and reports. The informal side is embodied in experience, tacit expertise, decision rationales, negotiated trade-offs, and local interpretations of uncertainty. In practice, project outcomes depend on the quality of the interaction between these two forms of knowledge more than on the mere existence of documentation.

The problem is that traditional approaches to project knowledge management have often privileged repositories over ecosystems. Knowledge has frequently been treated as something to be stored, archived, and retrieved, rather than as something continuously produced, interpreted, and transformed throughout the lifecycle of the project. The logic of "lessons learned at closure" is useful, but insufficient in environments where decisions must be made rapidly and where relevant knowledge emerges during planning, execution, escalation, exception handling, stakeholder interaction, and post-decision reflection. The spread of AI does not eliminate this challenge; it deepens it. By introducing digital agents capable of generating interpretations and suggestions, AI changes not only the speed of knowledge production but also the epistemic structure of project work.

It is here that the idea of a **Human–AI Hybrid Knowledge Ecosystem** becomes analytically useful. A hybrid knowledge ecosystem may be understood as a socio-technical arrangement in which knowledge assets are created, curated, circulated, validated, and applied through structured interactions among human actors, AI systems, project artifacts, governance rules, and sustainability criteria. The ecosystem perspective shifts the focus away from individual tools and toward the architecture of relationships among actors, assets, decisions, and controls. It foregrounds interdependence, traceability, and learning rather than isolated efficiency gains.

This approach is especially consistent with the INSPIRE PM view. The framework emphasizes that project environments must integrate multiple dimensions that are too often treated separately: innovation, sustainability, responsiveness, predictive capability, efficiency, and the intelligent use of technology. The same reasoning underlies PME, which frames projects as adaptive ecosystems where change is constant and where knowledge is continuously produced through feedback loops, interactions, and evolving constraints. In such a setting, knowledge cannot be reduced to documentation, nor AI reduced to automation. Instead, both must be treated as interdependent components of a wider knowledge ecology.

The significance of this shift is especially visible when sustainability is considered. Sustainable project management is not simply about adding ESG reporting to project closure. It is about embedding environmental, social, and governance considerations directly into decision parameters, planning choices, execution trade-offs, and accountability structures. If sustainability is to influence project outcomes meaningfully, it must be integrated into the knowledge ecosystem itself. It must appear in the criteria by which options are compared, in the indicators that are monitored, in the trade-offs that are documented, and in the decision rationales that are preserved.

3. INSPIRE PM as an Organizing Architecture for Human–AI Collaboration

The conceptual contribution of this paper lies in proposing INSPIRE PM as an organizing architecture for hybrid knowledge ecosystems in project-based organizations. This proposal does not imply that INSPIRE PM should function as a universal procedure. Rather, it suggests that the framework offers a robust conceptual scaffolding for aligning project knowledge, AI-enabled support, governance, and sustainability.

In the context of this study, INSPIRE PM performs at least three functions. First, it provides a way to organize the multiple dimensions that shape project knowledge in hybrid contexts: technology, sustainability, analytics, adaptability, efficiency, and innovation. Second, it supports the move from fragmented tool adoption to ecosystem design, that is, from asking "Which AI application should we use?" to asking "How should knowledge be produced, validated, and governed across the project lifecycle?" Third, it helps connect operational project decisions to broader organizational goals, including ESG commitments, resilience, and learning capacity.

This is particularly important because the adoption of AI in project environments often remains shallow. Organizations may introduce copilots, dashboards, or predictive models without rethinking the institutional logic within which such tools operate. They may generate content faster but still fail to establish who owns the knowledge, who validates it, who is accountable for decisions based on it, and how AI-supported outputs align with sustainability criteria. INSPIRE PM helps address this problem by situating technology within a broader managerial and governance logic rather than treating it as a standalone solution.

The human role remains central in this architecture. AI may support forecasting, classification, summarization, pattern recognition, or option generation, but the framing of the problem, the interpretation of context, the weighting of sustainability trade-offs, and the assumption of accountability remain human responsibilities. This principle is critical not only for ethical reasons but also for managerial ones. Projects are inherently contextual and relational. They involve ambiguity, competing stakeholder interests, constraints that change over time, and outcomes that cannot always be reduced to formal optimization. For this reason, hybrid intelligence in project management should not be understood as a substitution logic, but as an augmentation logic under governance.

4 Essential Knowledge Assets and Knowledge Flows Across the Project Lifecycle

The first research question concerns the knowledge assets and flows required to enable effective human–AI collaboration in sustainable project settings. Answering this question requires moving beyond a narrow focus on data and considering the broader ecology of project knowledge.

At minimum, a hybrid project knowledge ecosystem includes several categories of assets. The first category consists of **structured project artifacts**, such as schedules, requirements, work packages, budgets, change logs, and risk registers. These are the traditional backbone of project control and coordination. The second category includes **decision evidence**, such as assumptions, trade-off rationales, approvals, exceptions, and escalation records. These are often underdeveloped in practice, yet they are essential for accountability, learning, and auditability. The third category includes **contextual and tacit knowledge**, such as expert interpretation, stakeholder expectations, organizational memory, and situational judgment. The fourth category includes **ESG-related knowledge**, such as footprint indicators, compliance constraints, social impact considerations, and sustainability targets. Finally, in AI-enabled settings, there is also a fifth category: **AI-generated artifacts**, including summaries, predictions, recommendations, and classifications.

The critical issue is not only the existence of these assets, but the way they move across the lifecycle. During project initiation, knowledge is strongly strategic: business rationale, stakeholder expectations, sustainability commitments, and problem framing dominate. During planning, the ecosystem structures requirements, design assumptions, scenarios, success criteria, and governance choices. During execution and monitoring, the emphasis shifts toward operational signals, variance analysis, issue escalation, predictive warnings, exception handling, and real-time adaptation. During closure, the ecosystem should consolidate realized outcomes, assess whether sustainability targets were achieved, document decision rationales, and capture reusable learning for future initiatives.

AI can intervene at all these stages, but not in the same way. In initiation, AI may support horizon scanning, stakeholder pattern analysis, or synthesis of prior cases. In planning, it may assist in scenario comparison, requirements classification, or the drafting of structured documentation. In execution and monitoring, it may support anomaly detection, predictive alerts, and summarization of high-volume project communication. In closure, it may accelerate the synthesis of evidence and lessons. Yet the value of these interventions depends on whether the surrounding knowledge ecosystem can absorb them responsibly. AI output becomes useful only when linked to provenance, validation, decision rights, and feedback loops.

This observation also suggests a criterion for knowledge quality in hybrid ecosystems. Knowledge quality is not merely a function of correctness or completeness. In project contexts, it must also include timeliness, traceability, actionability, contextual relevance, and alignment with sustainability objectives. A technically correct output that cannot be trusted, traced, or connected to accountable decision-making is of limited value in a project-based organization.

5 Governance Mechanisms for Trustworthy AI-Augmented Project Knowledge

The second research question addresses governance. This is arguably the most critical dimension, because hybrid knowledge ecosystems can produce value only if they are trusted, and they can be trusted only if governance is explicit.

The first governance mechanism concerns **roles and decision rights**. Hybrid project environments require clarity regarding who owns each category of knowledge asset, who validates AI-generated content, which outputs remain advisory, and which decisions require mandatory human sign-off. If these boundaries are not explicit, organizations risk creating pseudo-automation in which AI-generated content silently acquires authority without corresponding accountability. The use of AI in project environments should therefore be bounded by role definitions that make human accountability visible rather than diffuse it.

A second governance mechanism concerns **provenance**. In project work, decisions often need to be justified ex post, whether for governance purposes, compliance reasons, stakeholder challenges, or organizational learning. When AI contributes to reports, forecasts, or recommendations, actors must be able to identify the origin of the underlying information, the nature of the model's contribution, and the extent to which human interpretation shaped the final output. Provenance is not simply a technical issue; it is a condition of organizational trust.

A third governance dimension concerns **privacy and ethical safeguards**. The abstract explicitly highlights privacy, bias awareness, and ethical constraints as central concerns in AI-enabled project work. This is especially important when project knowledge includes stakeholder data, sensitive commercial information, or ESG-relevant indicators that may have reputational and regulatory implications. Responsible governance requires that organizations distinguish clearly between permissible and impermissible use of data, define review thresholds for sensitive outputs, and maintain safeguards against automation bias and opaque decision-making.

A fourth mechanism concerns **human-in-the-loop review**. The phrase is often used loosely, but in the present framework it should be interpreted precisely. Human review is not satisfied merely because a person receives an AI output. It requires that a human actor has both the authority and the capability to assess the relevance, credibility, and implications of that output before it influences accountable project decisions. This includes the possibility of rejecting, revising, or re-framing the AI-supported suggestion. In other words, the human role must remain epistemically substantive, not ceremonially residual.

For these reasons, the governance of hybrid knowledge ecosystems should be light enough to remain usable, but explicit enough to remain meaningful. Over-bureaucratization would discourage adoption, while under-governance would undermine trust. The governance checklist proposed in this paper is intended precisely as a middle path: a practical mechanism for embedding provenance, accountability, privacy, and bias awareness into project knowledge practices without immobilizing the system.

6 Design Science Research and the Development of Actionable Artifacts

Methodologically, the study adopts a **Design Science Research** approach complemented by multiple-case evaluation, as anticipated in the abstract. This is appropriate because the objective is not only interpretive but also constructive: the research seeks to generate artifacts that can help organizations operationalize the concept of human–AI hybrid knowledge ecosystems in project settings.

The first artifact is an **INSPIRE-based Hybrid Knowledge Ecosystem Blueprint**. Its function is to make visible the architecture of the ecosystem: what knowledge exists, where it is produced, how it circulates, which actors validate it, how AI can intervene, and which controls apply. In practice, such a blueprint would help PMOs and project leaders move beyond vague digital ambition toward a structured operating model for project knowledge.

The second artifact is a **Decision Intelligence Loop** connecting predictive analytics, human judgment, decision logging, and feedback. The purpose of this artifact is to ensure that AI-supported pattern recognition and forecasting do not remain disconnected from actual managerial learning. Prediction alone does not create intelligence. Intelligence emerges when predictive signals are interpreted, linked to explicit criteria, translated into accountable decisions, and then compared with realized outcomes. The loop therefore turns project decision-making into a more traceable and learning-oriented process.

The third artifact is a **lightweight governance checklist for responsible GenAI use in project contexts**. This checklist is intended to address practical concerns such as provenance, privacy safeguards, bias awareness, and human review obligations. Its usefulness lies precisely in its minimalism: many organizations do not fail because they lack sophisticated governance theory, but because they lack operationally usable controls.

The evaluation strategy will involve two to three organizational cases in project-intensive contexts, combining interviews with PMO actors, project managers, data/AI roles, and sustainability leads, document analysis of project artifacts and decision logs, and validation workshops focused on feasibility, usefulness, and organizational fit. This methodological design is particularly suitable because hybrid knowledge ecosystems are inherently context-sensitive. Their form and value are likely to vary according to organizational maturity, regulatory exposure, digital capabilities, and sustainability priorities.

7 Measuring the Value of Hybrid Knowledge Ecosystems

The third research question concerns value measurement. If human–AI hybrid knowledge ecosystems are to be taken seriously as an organizational capability, their value must be assessed in ways that go beyond generic claims about efficiency.

At the level of **project performance**, relevant indicators may include decision cycle time, rework caused by poor information quality, speed of issue resolution, change-request quality, forecast reliability, stakeholder response time, and the quality of project coordination across functions. These indicators make visible whether the ecosystem improves the practical performance of project work rather than merely increasing the volume of digital activity.

At the level of **organizational resilience**, value should be assessed in terms of the ecosystem's capacity to sustain decision quality under uncertainty, adapt when assumptions change, preserve continuity when people or systems shift, and transfer learning from one project context to another. Resilience is especially relevant in project-based organizations because uncertainty is not episodic but structural. A hybrid ecosystem should therefore not only improve routine execution but also strengthen adaptive capacity.

At the level of **ESG outcomes**, the ecosystem should be evaluated according to its ability to make sustainability visible in decision processes. This implies not merely reporting environmental or social indicators after the fact, but incorporating them into the criteria used for planning, trade-off analysis, exception handling, and closure assessment. If a project claims to be sustainable but its ecosystem does not connect decisions to ESG evidence, then sustainability remains rhetorical rather than operational. INSPIRE PM is particularly valuable here because it explicitly positions sustainability as an integrated dimension of project practice rather than an external reporting add-on.

One important implication follows from this: the value of HA-HKE cannot be reduced to cost reduction or speed. It must be measured in a more balanced way, including trustworthiness, learning, traceability, adaptability, and sustainability alignment. This broader value logic is essential if AI is to contribute to genuinely sustainable project management rather than to short-term operational optimization alone.

8 Discussion

The framework proposed in this paper makes three conceptual moves that are significant for both research and practice. First, it reframes project management as a **knowledge ecosystem challenge** rather than only as a matter of methods, tools, or lifecycle control. This shift is necessary because the central problem in contemporary project environments is no longer information scarcity but knowledge fragmentation. Second, it positions AI as a participant in a socio-technical system rather than as an autonomous solution. This prevents the common mistake of interpreting AI outputs as self-validating. Third, it embeds sustainability directly into project knowledge architecture, thereby linking ESG concerns to daily project decisions and not only to post hoc evaluation.

This has important implications for PMOs. In a hybrid ecosystem logic, PMOs can no longer be understood only as custodians of templates, methods, or reporting cycles. They may increasingly need to act as stewards of knowledge architecture, responsible for designing how knowledge is structured, validated, and learned from across project portfolios. This also changes the capability profile required of project leaders. In addition to classical planning and coordination skills, hybrid project environments demand higher competence in evidence validation, AI literacy, model skepticism, knowledge traceability, and sustainability-aware decision framing.

At the same time, this perspective should not be romanticized. The presence of AI in the project environment does not guarantee better outcomes. Poorly designed ecosystems may generate new forms of opacity, shallow automation, misplaced trust, and compliance risk. The concept of HA-HKE is therefore valuable not because it assumes harmony between humans and AI, but because it foregrounds the tensions that must be governed:

speed versus traceability, automation versus accountability, scalability versus privacy, and predictive assistance versus contextual judgment.

9 Contributions, Limitations, and Future Research

The expected contributions of the study are threefold, in line with the extended abstract. First, the paper contributes **conceptually** by reframing project management as an intelligent knowledge ecosystem problem and by integrating socio-technical and sustainability dimensions under the INSPIRE PM architecture. Second, it contributes **methodologically** by proposing a design-oriented research path that translates abstract principles into practical artifacts. Third, it contributes **managerially** by offering PMOs and project leaders a concrete way to think about AI adoption not as isolated tooling, but as an issue of governance, learning, and organizational design.

The current limitation is that the paper remains, at this stage, primarily conceptual and design-oriented. The multiple-case evaluation planned in the abstract is therefore essential for testing whether the proposed artifacts are not only theoretically coherent but also operationally feasible and organizationally useful. Context will matter greatly: hybrid knowledge ecosystems are unlikely to take the same form in highly regulated sectors, sustainability-intensive settings, or digitally immature organizations.

Future research should therefore examine how different sectors, governance maturities, and sustainability priorities shape ecosystem adoption. It should also investigate which combinations of knowledge assets and AI capabilities are most valuable at different project lifecycle stages, how decision logs can effectively incorporate ESG trade-offs, and how organizations can avoid over-dependence on AI-generated recommendations while still benefiting from their analytical power.

10 Conclusion

This paper has argued that the challenge of AI in project-based organizations is not primarily one of access to tools, but one of knowledge ecosystem design. Building on INSPIRE PM and PME, it has proposed the concept of **Human–AI Hybrid Knowledge Ecosystems** as a way of understanding how project knowledge can be created, governed, and applied in organizations that seek both performance and sustainability.

The central claim is that sustainable project outcomes will increasingly depend on the ability to build ecosystems that are trustworthy, auditable, learning-oriented, and explicitly governed. In such ecosystems, AI may accelerate option generation, pattern detection, and predictive support, but humans remain responsible for framing problems, setting decision criteria, validating evidence, and taking accountable decisions. This is not a residual role; it is the condition under which hybrid intelligence remains organizationally legitimate.

INSPIRE PM offers a promising architecture for this transition because it makes it possible to connect innovation, technology, sustainability, responsiveness, predictive capability, and efficiency within a single project management logic. The next step is empirical validation through case-based evaluation, so that the conceptual model and the proposed design artifacts can be refined into deployable governance and operating models for real project-based organizations.

11 Ethics Declaration

This research did not require formal ethical clearance, as it is a conceptual and theoretical study involving no human participants, personal data collection, or processing of sensitive information.

12 AI Declaration

Generative AI tools were used as linguistic and stylistic support in preparing the paper. AI was not used to generate data, findings, or conclusions. The author reviewed and approved the final content and retains full responsibility for it.

References

Bassi, A. (2024). Scoprire il Project Management: lezioni e concetti essenziali (v.1). Unpublished manuscript.

Bassi, A. (2025). INSPIRE PM: Innovazione e sostenibilità per il Project Management del futuro (v.6). Unpublished manuscript.

Bassi, A. (2025). HPM: Hybrid Project Management – Abitare la tensione (v.1). PMWJ.

Davenport, T.H., & Prusak, L. (1998). Working Knowledge: How Organizations Manage What They Know. Harvard Business School Press.

ISO. (2020). ISO 21502:2020 Project, Programme and Portfolio Management — Guidance on Project Management. International Organization for Standardization.

Nonaka, I., & Takeuchi, H. (1995). The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation. Oxford University Press.

PMI. (2021). A Guide to the Project Management Body of Knowledge (PMBOK® Guide) – Seventh Edition. Project Management Institute.

Schiuma, G. (2011). Managing Knowledge Assets and Business Value Creation in Organizations: Measures and Dynamics. IGI Global.

AI-Driven Green Knowledge Management Processes and Startup Sustainable Performance: A Multidimensional Conceptual Framework

Furong Cai[1], Ettore Bolisani[1], Kathrin Kirchner[2], Tomas Cherkos Kassaneh[1], Enrico Scarso[1], Behrooz Moradi[1]

[1]Department of Management and Engineering, University of Padua, Vicenza, Italy

[2]Department of Engineering Technology, Technical University of Denmark, Ballerup, Denmark

Abstract

In the face of intensifying environmental challenges and rapid digital transformation, startups are increasingly expected to pursue sustainable development while navigating severe resource constraints and the inherent liability of newness typical of emerging ventures. Although existing literature acknowledges the critical role of Green Knowledge Management (GKM) in driving corporate sustainability, significant theoretical gaps remain. Prior studies frequently conceptualize GKM as a homogeneous construct, thereby masking the complex operational realities of its underlying dimensions. Furthermore, while Artificial Intelligence (AI) is rapidly altering organizational learning, limited academic attention has explored how AI specifically acts as a dynamic catalyst to enhance green knowledge processes within the entrepreneurial context. To address these critical omissions, this conceptual paper develops a comprehensive framework that elucidates how AI-driven GKM processes positively influence Sustainable Performance (SP) in startups. Drawing upon the Resource-Based View, the Knowledge-Based View, and the Dynamic Capabilities View, the proposed framework models GKM as a multidimensional construct comprising five distinct dimensions: green knowledge acquisition, storage, sharing, application, and creation. The model posits that robust GKM routines function as the core operational infrastructure enabling startups to translate environmental awareness into sustainability outcomes, encompassing long-term economic viability, reduced ecological footprints, and enhanced social responsibility. Crucially, this paper conceptualizes AI not as a generic technological solution, but as an advanced dynamic capability. AI is theorized to positively moderate the framework by acting as a digital catalyst that strengthens the relationship between GKM and SP. By providing analytical velocity and reducing prediction costs, AI empowers startups to overcome human capital limitations, rapidly sense subtle ecological shifts, mitigate trial-and-error risks, and efficiently apply internal knowledge assets to changing market demands. Integrating sustainability, knowledge management, and digital capabilities, this research advances theoretical discourse by deconstructing GKM into specific operational routines and refining the conceptual understanding of AI's moderating role. Ultimately, this conceptual study establishes a foundation for future empirical validation and offers actionable guidance for policymakers and founders to align knowledge routines with AI investments for enduring sustainable performance.

Keywords: Green Knowledge Management; Artificial Intelligence; Sustainable Performance; Startups; Dynamic Capabilities

Paper type: Academic Research Paper

1 Introduction

As global industries confront the urgent dual imperatives of digitalization and environmental sustainability, startups have emerged as vital engines of eco-innovation (Costa, 2024). However, unlike large incumbent corporations, startups operate in highly volatile environments characterized by severe resource constraints, high uncertainty, and the inherent liability of newness (Shehzad *et al.*, 2026; Sommer *et al.*, 2009). To survive and thrive amidst such turbulence, these emerging ventures must cultivate robust Knowledge Management (KM) capabilities. Traditionally, KM allows startups to systematically acquire, structure, and deploy intellectual capital to compensate for their critical lack of physical assets (Paoloni *et al.*, 2020). Yet, as the ecological mandate of this "twin transition" intensifies, general knowledge routines are no longer sufficient. Startups must evolve their learning architectures to process complex environmental data and embed ecological objectives into their core operations. Consequently, Green Knowledge Management (GKM) has emerged as a specialized and critical organizational asset, enabling startups to translate environmental awareness into tangible Sustainable Performance (SP) (Adusei, 2026; Nasir *et al.*, 2024).

Despite the acknowledged importance of KM in driving SP, the existing literature exhibits significant theoretical limitations. Primarily, prior studies often conceptualize GKM as a static, homogeneous construct. This monolithic perspective treats environmental knowledge routines as a theoretical black box, failing to capture the complex operational realities of its underlying processes. According to recent scale validations (Yu

et al., 2022), GKM is more accurately understood as a multidimensional construct comprising distinct sub-dimensions: green knowledge acquisition, storage, sharing, application, and creation. Treating these routines as a single variable obscures the specific operational bottlenecks and limits the precision of theoretical models predicting sustainable outcomes.

Furthermore, as the volume and complexity of ecological data continue to expand, traditional manual knowledge routines are becoming increasingly inadequate for startups. Artificial Intelligence (AI) has emerged as a transformative digital capability, fundamentally altering how organizations process information and generate insights (Oyekunle and Boohene, 2024). Recent scholarship explores the nexus of AI, KM, and sustainable outcomes. For instance, Al-Balushi et al. (2025) demonstrated that AI adoption significantly drives green innovation, highlighting specific knowledge practices like green knowledge sharing as critical boundary conditions. Similarly, emerging conceptual frameworks (Cai *et al.*, 2025) suggest that AI and big data analytics serve as crucial digital enablers for startups, strengthening the impact of GKM by accelerating sensing and learning processes. However, these foundational studies tend to either isolate a single facet of KM or treat GKM as an aggregated, macro-level variable. Limited academic attention has explored how AI specifically acts as a dynamic catalyst to enhance the complete, multidimensional pipeline of GKM within the entrepreneurial context. While often labeled as a generic technological solution, the literature lacks a coherent framework articulating how AI-driven capabilities specifically empower startups to accelerate their distinct routines of acquiring, storing, sharing, applying, and creating green knowledge to drive SP.

To address these critical research gaps, this study develops a comprehensive framework that elucidates how AI-driven GKM processes directly drive SP in startups. Drawing upon the Resource-Based View (RBV), the Knowledge-Based View (KBV), and the Dynamic Capabilities View (DCV), this study disaggregates GKM into its five foundational dimensions to map specific pathways of eco-innovation. The proposed model posits multidimensional GKM as the core operational mechanism for achieving sustainability outcomes, while conceptualizing AI as a potent, dynamic enabling capability that positively moderates and strengthens these knowledge-based relationships. Consequently, this paper offers three primary contributions. First, it advances KM theory by unpacking GKM into heterogeneous processes, moving beyond the monolithic approach. Second, it refines the conceptual understanding of AI by theorizing its role as a dynamic digital catalyst tailored to enhance knowledge mechanisms. Third, it provides a robust, proposition-driven theoretical foundation for future empirical research on AI-augmented knowledge processes in sustainable startups.

The remainder of this paper proceeds as follows: Section 2 establishes the theoretical background and defines the core constructs. Section 3 develops the conceptual framework and theoretical propositions. Section 4 discusses theoretical contributions and practical implications. Finally, Section 5 concludes by outlining limitations and avenues for future empirical research.

2 Theoretical Background

2.1 Theoretical foundations: integrating RBV, KBV, and DCV

To conceptualize how startups achieve sustainable outcomes in highly volatile environments, this framework integrates three foundational strategic management perspectives. The RBV (Barney, 1991) postulates that organizations achieve competitive advantages by accumulating tangible and intangible resources that are valuable, rare, inimitable, and non-substitutable. Building directly upon this resource-centric logic, the KBV (Grant, 1996) argues that in complex, high-technology environments, knowledge is the most strategically significant asset a firm can possess. In the context of the twin transition, specialized green knowledge serves as this ultimate strategic asset. Recent systematic reviews affirm that these knowledge and resource perspectives remain the dominant theoretical lenses for understanding how GKM translates into sustainable business performance (Adusei, 2026).

However, the static nature of traditional resource theories renders them insufficient for startups operating in rapidly evolving digital and ecological landscapes. To bridge this gap, the DCV (Teece, 2007) extends these perspectives by emphasizing a firm's dynamic capability to systematically sense, seize, and reconfigure internal and external competences. Through this integrated theoretical lens, this paper establishes a highly complementary dynamic: AI acts as the ultimate digital sensing mechanism providing analytical velocity, while GKM functions as the dynamic operational mechanism responsible for the actual seizing and reconfiguration of organizational resources.

2.2 Green Knowledge Management (GKM) as a multidimensional mechanism

To address this theoretical limitation, this framework draws upon Knowledge Lifecycle Theory (Alavi and Leidner, 2001; Dalkir, 2013), which posits that KM is not a single action but a complex, sequential pipeline of distinct cognitive and social routines. According to this lifecycle perspective, treating green knowledge as a homogeneous entity obscures the specific capability bottlenecks a firm might face. A startup might excel at acquiring external environmental data but fail completely at sharing it internally or applying it to production. Therefore, to accurately map capability development, this framework adopts the multidimensional conceptualization validated by Yu et al. (2022) and strongly corroborated by recent comprehensive GKM empirical models (Abbas and Khan, 2023; Widyanti *et al.*, 2024). Specifically, GKM is defined as a multidimensional construct comprising five distinct, interrelated sub-dimensions that must function as an uninterrupted pipeline to achieve sustainable outcomes:

- Green Knowledge Acquisition: The capability to proactively gather and sense external environmental information, clean technology updates, and stakeholder requirements.
- Green Knowledge Storage: The infrastructural capacity to codify, structure, and securely retain ecological expertise within the organizational memory, preventing knowledge loss due to high employee turnover typical in startups.
- Green Knowledge Sharing: The internal socio-technical mechanisms that facilitate the horizontal diffusion of environmental best practices, fostering a collaborative climate where ecological insights are seamlessly exchanged across human and digital subsystems.
- Green Knowledge Application: The practical execution phase where stored and shared insights are actively mobilized to improve eco-friendly operations, ensure regulatory compliance, and inform sustainable business strategies.
- Green Knowledge Creation: The highest-order cognitive capability, building upon the SECI model of knowledge conversion (Nonaka, 1994), which involves generating novel, proprietary environmental solutions and original eco-innovations.

2.3 Artificial Intelligence (AI) as a dynamic digital catalyst

A central theoretical contribution of this framework is the conceptual elevation of AI. Rather than viewing AI merely as a generic information technology tool, this paper conceptualizes AI as an advanced organizational capability and a dynamic digital catalyst. Building upon foundational theories of AI capability (Mikalef and Gupta, 2021) and the automation-augmentation paradigm (Raisch and Krakowski, 2021), AI represents a transformative cognitive infrastructure. It fundamentally alters an organization's boundaries by not only automating routine technical tasks but also deeply augmenting human intelligence.

In the startup context, these AI capabilities encompass the deployment of statistical, self-learning, and predictive algorithms across various operational levels. By leveraging machine learning, AI substantially reduces the prediction costs associated with highly complex and volatile ecological environments (Agrawal *et al.*, 2022). Through a dynamic capability lens, AI profoundly enhances a startup's ability to process massive ecological datasets in real time. It automates complex pattern recognition and informs shared decision-making, providing the analytical velocity required to tackle sustainability challenges. However, AI alone lacks contextual judgment. By acting as an advanced cognitive extension, AI mitigates the severe human capital constraints typical of emerging ventures, yet it fundamentally relies on the human-centric governance of GKM to translate its algorithmic predictions into trusted, sustainable actions.

2.4 Sustainable Performance (SP)

The ultimate dependent variable in this framework is SP. Moving beyond narrow economic metrics that focus solely on short-term profitability, SP is conceptualized as a multidimensional construct rooted in the Triple Bottom Line approach (Elkington, 1998) and the principles of corporate sustainable development (Bansal, 2005).

Aligned with the comprehensive metrics utilized by Khan et al. (2024), SP in startups is evaluated across three interconnected pillars. Economically, it reflects competitive advantages in cost-saving, operational efficiency, and sustained value creation. Environmentally, it signifies the successful implementation of initiatives to reduce, reuse, and recycle, thereby minimizing the negative ecological footprint of the firm's operations and products. Socially, it encompasses organizational resilience, respect for business ethics, and the pursuit of being a responsible corporate citizen. This multidimensional conceptualization ensures that the outcomes of AI-driven green knowledge routines are evaluated comprehensively, capturing both immediate market viability and long-term ecological impact.

3 Conceptual Framework and Propositions

This section develops a comprehensive conceptual framework delineating the direct and moderating pathways linking GKM, AI, and SP within the specific context of emerging startups.

3.1 The impact of GKM on SP

To achieve SP, startups must move beyond superficial environmental compliance and embed ecological considerations deep within their core organizational routines. According to the knowledge lifecycle perspective, this embedding process critically depends on a robust GKM pipeline. Recent systematic reviews demonstrate that effective green knowledge routines act as independent predictors for corporate sustainability (Adusei, 2026). Specifically, scholars highlight that GKM practices directly enhance sustainable development by fostering green technological innovation, minimizing ecological footprints, and optimizing resource efficiency (Khan *et al.*, 2024; Wang *et al.*, 2022). By systematically managing environmental knowledge, organizations successfully transform green intellectual capital into tangible economic and socio-environmental value (Zhang *et al.*, 2025). Theoretical models increasingly treat GKM not as a singular action but as an integrated system of sequential processes that translates ecological intent into actionable sustainable outcomes (Widyanti *et al.*, 2024). Therefore, contextualizing these mechanisms for emerging ventures, this study posits:

Proposition 1 (P1): *GKM composed of acquisition, storage, sharing, application, and creation, acts as a comprehensive driver positively influencing the SP of startups.*

To unpack this overarching relationship, it is theoretically imperative to examine the distinct operational routines that constitute this pipeline. Because environmental challenges are inherently complex and systemic, the survival of startups in the twin transition depends fundamentally on their boundary-spanning capabilities (De Marchi, 2012). Proactively acquiring external environmental intelligence enables startups to accurately sense rapid regulatory shifts and identify emerging clean technologies before their competitors (Albort-Morant *et al.*, 2016). This early sensing capability serves as a critical defensive mechanism that minimizes institutional risks while simultaneously uncovering lucrative green market niches, directly enhancing long-term economic resilience (Xue *et al.*, 2019). Recent empirical studies validate this mechanism, demonstrating that green knowledge acquisition serves as a critical strategic lever. Specifically, robust knowledge acquisition activities are essential for translating green technological innovation into superior environmental performance (Sahoo *et al.*, 2023). Furthermore, acquiring green knowledge acts as a vital operational mechanism enabling smaller, resource-constrained ventures to successfully convert their strategic environmental intent into SP (Baquero, 2025). Therefore, the ability for startups to rapidly internalize external ecological data is a prerequisite for subsequent knowledge routines and overall sustainability.

Proposition 1a (P1a): *Green knowledge acquisition positively influences the SP of startups.*

Once knowledge is acquired, retention becomes the next critical hurdle. Startups inherently suffer from high employee turnover and the liability of newness. Without structural storage, tacit ecological expertise leaves when key employees depart. Robust digital storage routines convert fragmented human capital into formalized organizational memory. Recent literature emphasizes that the structural preservation of environmental knowledge through the entire knowledge lifecycle is indispensable for maintaining operational continuity and achieving long-term sustainable development (Widyanti *et al.*, 2024). Furthermore, utilizing appropriate technological infrastructure to securely retain green knowledge prevents severe knowledge attrition and creates a reliable database for continuous sustainability improvements (Al-Faouri, 2023; Kordab *et al.*, 2020), safeguarding startups against their liability of newness.

Proposition 1b (P1b): *Green knowledge storage positively influences the SP of startups.*

Furthermore, complex environmental challenges cannot be solved in silos. Internal socio-technical sharing mechanisms foster a collaborative climate that rapidly diffuses green best practices across different functional teams. This horizontal diffusion translates implicit individual awareness into a collective organizational green culture. Recent empirical studies strongly underscore the criticality of this dimension. Specifically within resource-constrained organizational settings, active green knowledge sharing functions as a pivotal mechanism that translates green innovation strategies into tangible sustainability outcomes (Rahman *et al.*, 2026). Moreover, establishing robust internal sharing mechanisms directly drives sustainable business performance and secures long-term competitive advantage (Zairbani and Senthil Kumar, 2025). Facilitating seamless communication across agile teams reduces operational waste and builds a unified entrepreneurial strategy vital for small enterprises' survival (Junejo *et al.*, 2025).

Proposition 1c (P1c): *Green knowledge sharing positively influences the SP of startups.*

Crucially, the materialization of SP relies heavily on execution. Acquired and stored intelligence remains a latent asset until it is practically applied. The knowledge application phase ensures that ecological data is actively

mobilized to optimize energy efficiency, reduce material footprints, and streamline supply chains. Executing these eco-friendly operational adjustments directly mitigates negative environmental impacts while simultaneously strengthening the foundational capabilities required for green innovation (Abbas and Khan, 2023). Consequently, translating theoretical environmental awareness into applied practices is the definitive step for securing comprehensive sustainability in highly dynamic environments where startups operate with limited resources (Rahoo *et al.*, 2026).

Proposition 1d (P1d): *Green knowledge application positively influences the SP of startups.*

Finally, as the highest-order cognitive capability within this multidimensional pipeline, green knowledge creation enables startups to transcend mere regulatory compliance and generate proprietary eco-innovations. By synthesizing and recombining existing environmental insights, startups can develop novel green products, circular business models, and distinct operational paradigms. Recent empirical evidence specifically targeting emerging enterprises confirms that the continuous creation of fresh green knowledge is a decisive factor for achieving sustainable development, particularly when leveraging digital capabilities (Zhuge *et al.*, 2023). Furthermore, generating original ecological solutions allows these ventures to capture premium eco-conscious market segments and establish an enduring green competitive advantage in the modern industrial landscape (Abbas and Khan, 2023). Therefore, the ultimate realization of long-term sustainability hinges on a startup's capacity to continuously invent and deploy novel environmental knowledge.

Proposition 1e (P1e): *Green knowledge creation positively influences the SP of startups.*

3.2 The moderating role of AI as a dynamic catalyst

While multidimensional GKM provides the necessary operational pipeline for sustainability, startups are fundamentally bottlenecked by severe human capital limitations and financial resource constraints. They often lack the sheer manpower required to process massive volumes of ecological data or the financial buffer to absorb the high trial-and-error costs associated with traditional green innovation. Therefore, the actual conversion of green knowledge into SP requires a powerful moderating force to overcome these inherent liabilities.

Operating as an advanced dynamic capability, AI provides this moderating force by fundamentally altering the speed, scale, and efficiency of organizational learning. Recent empirical evidence strongly validates this catalytic function within resource-constrained environments. Specifically, integrating AI into smaller enterprises significantly amplifies the positive impact of green knowledge sharing and dynamic capabilities on sustainable outcomes (Rahman *et al.*, 2026). Furthermore, AI acts as a critical technological boundary condition that strengthens the mechanisms through which organizations convert overall GKM routines into actionable green intellectual capital (Khan *et al.*, 2024; Zhang *et al.*, 2025). When a startup possesses robust GKM routines, AI functions as a digital accelerator for the entire knowledge pipeline. For instance, machine learning algorithms can automate the continuous tracking of global environmental regulations, substantially reducing the time and cost of knowledge acquisition. Similarly, predictive analytics can instantly translate stored ecological data into real-time operational adjustments, maximizing the efficiency of knowledge application.

Furthermore, AI bridges the critical gap between knowledge possession and value realization. In fast-paced startup ecosystems, merely possessing green knowledge does not automatically guarantee SP if that knowledge cannot be deployed rapidly. By utilizing advanced algorithms to simulate eco-friendly product designs and optimize resource allocation, AI significantly reduces the time-to-market for green innovations. Consequently, AI empowers emerging ventures to overcome their size-related disadvantages, ensuring that their internal green knowledge routines are translated into optimal economic and environmental returns at a highly accelerated pace.

Proposition 2 (P2): *AI acts as a dynamic catalyst that positively moderates the relationship between GKM and SP, amplifying the impact of green knowledge routines on sustainable outcomes.*

3.3 Summary of the conceptual model

Synthesizing the theoretical derivations and the propositions outlined above, this study proposes a comprehensive conceptual framework, as illustrated in Figure 1. The model maps the critical architecture through which startups navigate the twin transition despite severe resource and human capital constraints. It positions the five distinct sub-dimensions of GKM as the core operational pipeline driving SP. Crucially, AI is integrated as a dynamic digital catalyst that positively moderates this central relationship, accelerating the conversion of ecological knowledge into tangible economic and socio-environmental value.

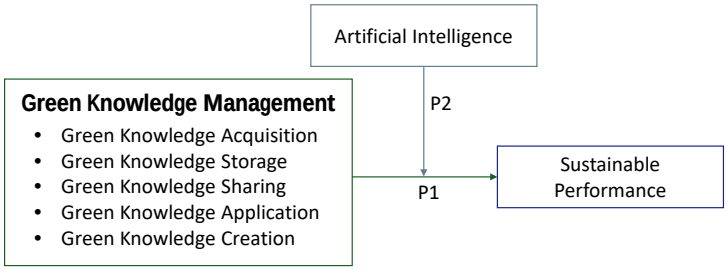


Figure 1: The conceptual framework of AI-driven GKM in startups

4 Discussion

4.1 Synthesizing the core framework

This study set out to address a critical paradox in the modern entrepreneurial landscape: how startups can successfully navigate the highly demanding twin transition of digitalization and sustainability. By integrating the RBV, KBV, and DCV, the developed framework illustrates that sustainable outcomes are not generated by sheer technological adoption alone. Instead, they require a meticulously orchestrated knowledge pipeline. The propositions advanced in this study highlight that GKM acts as the foundational operational infrastructure for sustainable performance, while AI serves as the dynamic digital catalyst that accelerates the flow of intelligence through this system.

4.2 Theoretical contributions

This research advances the intersection of sustainable entrepreneurship, KM, and AI through two primary theoretical contributions. First, it challenges the prevailing monolithic treatment of GKM by leveraging knowledge lifecycle theory to deconstruct it into a five-stage, sequential pipeline: acquisition, storage, sharing, application, and creation. This granular lens establishes that achieving SP relies on the uninterrupted flow of intelligence, where a failure in any single dimension can severely bottleneck the organizational eco-innovation process. Second, the framework refines the conceptualization of AI within the sustainability literature. Rather than viewing AI as a deterministic or universal performance driver, this study positions it as an advanced dynamic capability. This nuanced perspective moves the academic conversation beyond simple technology adoption metrics, theorizing that the true strategic value of AI lies in its ability to catalytically accelerate and optimize these underlying human-centric green knowledge architectures.

4.3 Practical implications

For practitioners navigating the twin transition, this framework offers valuable strategic perspectives across both organizational and institutional levels. Primarily, it cautions startup founders against the unstructured procurement of AI, emphasizing that predictive algorithms generally require a structurally sound data pipeline to deliver meaningful sustainability outcomes. Consequently, resource-constrained ventures are advised to prioritize the foundational routines of acquiring and securely storing green data before expecting digital tools to spontaneously drive SP. Furthermore, while AI significantly accelerates data processing, founders should actively cultivate a collaborative culture, recognizing that human intuition remains essential for sharing complex ecological insights and generating original eco-innovations. At the institutional level, policymakers and public incubators could consider shifting from generic technological grants toward targeted subsidies. By specifically funding AI-augmented storage infrastructures and environmental data-sharing platforms, institutional support can help mitigate capability bottlenecks, assisting startups in transforming stringent regulatory compliance pressures into genuine green innovation capabilities.

5 Conclusions and Future Research

The pursuit of SP in the digital era requires startups to move beyond superficial green posturing and meticulously orchestrate an AI-augmented knowledge pipeline. However, as a theoretical exposition, this paper acknowledges inherent limitations that simultaneously provide clear pathways for future empirical research. Primarily, the absence of quantitative testing means the proposed framework serves as a foundational blueprint requiring rigorous validation. Future researchers are strongly encouraged to empirically test this multidimensional model using advanced statistical techniques, such as Partial Least Squares Structural Equation

Modeling (PLS-SEM). By collecting large-scale, cross-sectional survey data from startups operating in diverse institutional contexts, scholars can quantify the specific path coefficients linking the five GKM dimensions to sustainable outcomes. Crucially, subsequent studies should build upon this framework by exploring the strategic intent or high-level organizational antecedents that drive the initial adoption of these knowledge pipelines. Furthermore, sophisticated moderation analyses should be conducted to investigate whether the accelerating power of AI varies in intensity across the different cognitive phases of the knowledge lifecycle, specifically testing if explicit knowledge routines benefit more from digital catalysts than highly tacit human processes.

6 Acknowledgements

This work has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 101034319 and from the European Union - NextGenerationEU. We also thank the anonymous reviewers for their constructive feedback, which has significantly improved the quality of this manuscript.

7 Ethics Declaration

No ethical approval was required as the study is conceptual in nature and does not involve empirical research.

8 AI Declaration

The authors used Gemini solely for language refinement during the preparation of this paper. The tool provided grammar, phrasing, and proofreading suggestions, which the authors reviewed and incorporated into the manuscript.

References

Abbas, J. and Khan, S.M. (2023) "Green Knowledge Management and Organizational Green Culture: An Interaction for Organizational Green Innovation and Green Performance", *Journal of Knowledge Management*, Vol. 27, No. 7, pp. 1852–1870.

Adusei, E. (2026) "Green Knowledge Management Practices and Sustainable Business Performance: Strategic Insights from Past Research and Future Directions", *Knowledge and Process Management*.

Agrawal, A., Gans, J. and Goldfarb, A. (2022) *Prediction Machines, Updated and Expanded: The Simple Economics of Artificial Intelligence*, Harvard Business Press, Boston.

Alavi, M. and Leidner, D.E. (2001) "Review: Knowledge Management and Knowledge Management Systems: Conceptual Foundations and Research Issues", *MIS Quarterly*, Vol. 25, No. 1, pp. 107–136.

Al-Balushi, H.A., Singh, H. and Saleem, I. (2025) "Unlocking Sustainable Performance in the Health-Care Sector: The Dynamic Nexus of Artificial Intelligence, Green Innovation and Green Knowledge Sharing", *Society and Business Review*, Vol. 20, No. 3, pp. 545–565.

Albort-Morant, G., Leal-Millán, A. and Cepeda-Carrión, G. (2016) "The Antecedents of Green Innovation Performance: A Model of Learning and Capabilities", *Journal of Business Research*, Vol. 69, No. 11, pp. 4912–4917.

Al-Faouri, A.H. (2023) "Green Knowledge Management and Technology for Organizational Sustainability: The Mediating Role of Knowledge-Based Leadership", *Cogent Business & Management*, Vol. 10, No. 3, p. 2262694.

Bansal, P. (2005) "Evolving Sustainably: A Longitudinal Study of Corporate Sustainable Development", *Strategic Management Journal*, Vol. 26, No. 3, pp. 197–218.

Baquero, A. (2025) "Optimizing Green Knowledge Acquisition Through Entrepreneurial Orientation and Resource Orchestration for Sustainable Business Performance", *Marketing Intelligence & Planning*, Vol. 43, No. 2, pp. 241–271.

Barney, J. (1991) "Firm Resources and Sustained Competitive Advantage", *Journal of Management*, Vol. 17, No. 1, pp. 99-120.

Cai, F. et al. (2025) "How Does AI-Driven Knowledge Management Enhance Sustainability of Startups? A Conceptual Framework", *European Conference on Knowledge Management*, Vol. 26, No. 1, pp. 173–182.

Costa, A. (2024) *Navigating the Twin Transition: Digital Transformation and Sustainability in Low-Tech Industries: A Small and Medium-Sized Companies' Perspective*, TAB edizioni, Rome.

Dalkir, K. (2013) *Knowledge Management in Theory and Practice*, Routledge, London.

De Marchi, V. (2012) "Environmental Innovation and R&D Cooperation: Empirical Evidence from Spanish Manufacturing Firms", *Research Policy*, Vol. 41, No. 3, pp. 614–623.

Elkington, J. (1998) *Cannibals with Forks: The Triple Bottom Line of 21st Century Business*, New Society Publishers, Stony Creek.

Grant, R.M. (1996) "Toward a Knowledge-Based Theory of the Firm", *Strategic Management Journal*, Vol. 17, No. S2, pp. 109–122.

Junejo, I. et al. (2025) “Sustainable Performance in Manufacturing SMEs of Developing Country Through Green Supply Chain Management and Green Knowledge Management: A Mediated Moderation Model”, *Environmental Engineering and Management Journal*, Vol. 24, No. 9, pp. 1969–1985.

Khan, A.N., Mehmood, K. and Kwan, H.K. (2024) “Green Knowledge Management: A Key Driver of Green Technology Innovation and Sustainable Performance in the Construction Organizations”, *Journal of Innovation & Knowledge*, Vol. 9, No. 1, p. 100455.

Kordab, M., Raudeliūnienė, J. and Meidutė-Kavaliauskienė, I. (2020) “Mediating Role of Knowledge Management in the Relationship Between Organizational Learning and Sustainable Organizational Performance”, *Sustainability*, Vol. 12, No. 23, p. 10061.

Mikalef, P. and Gupta, M. (2021) “Artificial Intelligence Capability: Conceptualization, Measurement Calibration, and Empirical Study on Its Impact on Organizational Creativity and Firm Performance”, *Information & Management*, Vol. 58, No. 3, p. 103434.

Nasir, A. et al. (2024) “The Impact of Knowledge Management on Sustainable Performance with Mediation Effect of Green Innovation in Malaysian SMEs”, *VINE Journal of Information and Knowledge Management Systems*.

Nonaka, I. (1994) “A Dynamic Theory of Organizational Knowledge Creation”, *Organization Science*, Vol. 5, No. 1, pp. 14–37.

Oyekunle, D. and Boohene, D. (2024) “Digital Transformation Potential: The Role of Artificial Intelligence in Business”, *International Journal of Professional Business Review*, Vol. 9, No. 3, p. 1.

Paoloni, M. et al. (2020) “Knowledge Management, Intellectual Capital and Entrepreneurship: A Structured Literature Review”, *Journal of Knowledge Management*, Vol. 24, No. 8, pp. 1797–1818.

Rahman, S. et al. (2026) “Power of Green Capabilities and Artificial Intelligence (AI): Understanding How and When Green Innovation Promotes Sustainability”, *Business Strategy and the Environment*, Vol. 35, No. 2, pp. 2502–2525.

Rahoo, L.A. et al. (2026) “Impact of Green Knowledge Management on Organizational Sustainability in Higher Education: The Mediating Role of Knowledge-Oriented Leadership”, *Center for Management Science Research*, Vol. 4, No. 1.

Raisch, S. and Krakowski, S. (2021) “Artificial Intelligence and Management: The Automation–Augmentation Paradox”, *Academy of Management Review*, Vol. 46, No. 1, pp. 192–210.

Sahoo, S., Kumar, A. and Upadhyay, A. (2023) “How Do Green Knowledge Management and Green Technology Innovation Impact Corporate Environmental Performance? Understanding the Role of Green Knowledge Acquisition”, *Business Strategy and the Environment*, Vol. 32, No. 1, pp. 551–569.

Shehzad, N. et al. (2026) “Navigating Uncertainty: Conceptualizing Resilience in Startups Through a Systematic Review”, *International Entrepreneurship and Management Journal*, Vol. 22, No. 2, p. 61.

Sommer, S.C., Loch, C.H. and Dong, J. (2009) “Managing Complexity and Unforeseeable Uncertainty in Startup Companies: An Empirical Study”, *Organization Science*, Vol. 20, No. 1, pp. 118–133.

Teece, D.J. (2007) “Explicating Dynamic Capabilities: The Nature and Microfoundations of (Sustainable) Enterprise Performance”, *Strategic Management Journal*, Vol. 28, No. 13, pp. 1319–1350.

Wang, S. et al. (2022) “Achieving Green Innovation and Sustainable Development Goals Through Green Knowledge Management: Moderating Role of Organizational Green Culture”, *Journal of Innovation & Knowledge*, Vol. 7, No. 4, p. 100272.

Widyanti, R., Rajiani, I. and Basuki, B. (2024) “Green Knowledge Management to Achieve Corporate Sustainable Development”, *Journal of Infrastructure, Policy and Development*, Vol. 8, No. 2, p. 2844.

Xue, M., Boadu, F. and Xie, Y. (2019) “The Penetration of Green Innovation on Firm Performance: Effects of Absorptive Capacity and Managerial Environmental Concern”, *Sustainability*, Vol. 11, No. 9, p. 24555.

Yu, S. et al. (2022) “Green Knowledge Management: Scale Development and Validation”, *Journal of Innovation & Knowledge*, Vol. 7, No. 4, p. 100244.

Zairbani, A. and Senthil Kumar, J.P. (2025) “How Does Green Knowledge Sharing Improve Sustainable Business Performance?”, *Knowledge and Process Management*, Vol. 32, No. 1, pp. 42–53.

Zhang, J. et al. (2025) “Empowering Eco-Innovation: How Artificial Intelligence and Green Leadership Enhance Knowledge Capital for Sustainable Performance”, *Journal of Environmental Management*, Vol. 394, p. 127145.

Zhuge, K. et al. (2023) “Does Digital Capability Promote Sustainable Development of New Ventures? The Dual Impact of Green Knowledge Creation and Green Pressure”, *International Journal of Environmental Research and Public Health*, Vol. 20, No. 3, p. 2274.

Digital ESG Communication in Sustainable Fashion Consumption: Comparing Italy and South Korea

Strazzullo Serena[1], Ianniello Sara[1], Lee Jiyeon[2], Moon Ilkyeong[2]
[1]University of Naples Federico II, Naples, Italy
[2]Seoul National University, Seoul, Korea

Abstract

The fashion industry is increasingly challenged to demonstrate credible sustainability commitments in response to its significant environmental and social impacts, globalised supply chains, and persistent concerns over unsustainable production and consumption patterns. In this context, Environmental, Social, and Governance (ESG) communication has become a central tool by which fashion companies disclose their sustainability efforts and engage consumers. However, the effectiveness of ESG communication in influencing consumer behaviour remains uncertain. This is particularly relevant in digital environments, where sustainability claims are often difficult to verify and where information asymmetry, consumer scepticism, and greenwashing concerns may weaken the credibility of corporate messages. Moreover, consumers' interpretation of sustainability communication may differ across geographical contexts. Hence, the study adopts a narrative literature review approach, focusing on two main streams of research: ESG communication practices in the fashion industry and consumer behaviour studies related to sustainable consumption, with specific attention to Italy and South Korea, two relevant and contrasting contexts.

Building on signalling theory, the study develops an empirical model that integrates the insights emerging from the narrative review with cross-country evidence. In the proposed model, ESG communication practices are conceptualised as sustainability signals. Sustainability predisposition and digital and social influence are introduced as antecedents, shaping consumers' interpretation of these signals. In line with the theory, the interpretation of sustainability signals is expected to influence consumers' behavioural intention, which is linked to actual purchase behaviour, while behavioural barriers are considered as factors that may affect this relation.

In doing so, the study advances a model that can be empirically tested through multi-group analysis to assess whether the proposed relationships operate differently among Italian and South Korean consumers. Generally, the study offers a theoretical and practical foundation for understanding how digital ESG communication becomes meaningful, credible, and behaviourally relevant in different contexts.

Keywords: ESG, digital technologies, fashion industry, trust.

Paper type: Academic Research Paper

1 Introduction

The fashion industry is widely recognised as one of the most environmentally and socially impactful sectors, due to its intensive resource consumption, globalised supply chains, and fast production cycles that generate significant environmental and social externalities (Ronda, 2024). In response to increasing regulatory pressure and societal expectations, firms have gradually integrated Environmental, Social, and Governance (ESG) practices into their strategies, positioning sustainability as a source of competitive advantage and differentiation (Hörisch et al., 2020). At the same time, consumers have become more aware of environmental and ethical issues, showing growing interest in the sustainability performance of fashion products and brands (Sharma et al., 2020). However, despite this increasing awareness, sustainable consumption in fashion remains inconsistent. Consumers often fail to translate pro-environmental attitudes into actual purchasing behaviour, giving rise to the well-documented attitude-behaviour gap (Annunziata et al., 2023; Park & Lin, 2020). This conflict is particularly evident in fashion, where purchasing decisions are influenced by cognitive evaluations, and hedonic, social, and symbolic factors, as well as by habits and price sensitivity (Colasante et al., 2025). As a result, even when consumers express positive attitudes towards sustainability, these do not necessarily lead to consistent behavioural outcomes.

A particularly relevant but still underexplored issue concerns how firms communicate ESG efforts in digital environments and how such communication translates into consumer trust.

From a signalling perspective (Williams & Hodges, 2022), sustainability communication can reduce information asymmetry by making sustainability attributes more observable and interpretable (Fu et al., 2023). However, the effectiveness of these signals depends on their perceived credibility and transparency. Trust literature complements this view by highlighting that consumers respond positively only when information is perceived as reliable, especially in contexts characterised by uncertainty and risk (Tan et al., 2025). In the

absence of credible signals, sustainability claims may generate scepticism and weaken consumer responses, especially in fashion contexts typically exposed to greenwashing (Tan et al., 2025). While prior research has demonstrated that perceived transparency of signals enhances purchasing intention, this mechanism has not been systematically integrated into the study of ESG communication in the fashion domain.

This gap becomes even more relevant when considering cross-country differences (Koszewska et al., 2020; Whang et al., 2015). More broadly, recent review evidence on ESG and consumer behaviour highlights a strong geographical and contextual concentration of findings, calling for a broader consideration of socio-cultural environments in explaining consumer responses (Tripopsakul & Puriwat, 2022). Consumer responses to sustainability cues are not uniform across contexts, as they are shaped by socio-cultural factors, market structures, consumption norms (Whang et al., 2015), shopping habits, and consumption patterns (Colasante et al., 2025). Evidence suggests that different markets may prioritise distinct dimensions of sustainability (Hong & Kang, 2019), such as ethical concerns, economic considerations, or symbolic value, leading to heterogeneous interpretations of ESG communication (Tripopsakul & Puriwat, 2022). This means that consumers in different contexts may also differ in their ability and willingness to process such information. For instance, limited awareness of ESG practices and difficulties in accessing or interpreting firm disclosures can generate significant information resistance, weakening the translation of sustainability preferences into actual purchasing behaviour (Leonelli et al., 2025).

Building on these premises, this study addresses the need to analyse digital ESG communication practices and their influence on purchasing behaviour by comparing Italy and South Korea. Specifically, the analysis builds on a narrative literature review aimed at identifying the most relevant contributions to evaluate contextual differences and relationships linking sustainable fashion consumption with digital ESG communication practices.

Finally, the research proposes an empirical model, grounded in signalling theory, in which digital ESG communication acts as a signal, and consumer interpretation constitutes the main relational pathway leading to purchasing behaviour.

2 Theoretical Background: Signalling Theory in Sustainable Fashion Consumption

The study of sustainable fashion consumption is characterised by a fundamental condition of information asymmetry between firms and consumers. Sustainability attributes, such as ethical sourcing, environmental impact, and labour practices, are not directly observable at the point of purchase and therefore belong to the category of credence attributes. As a result, consumers are unable to independently verify the authenticity of sustainability claims and must rely on external information provided by firms. Within this context, signalling theory has increasingly been adopted to explain how firms communicate unobservable product and process characteristics (Market et al., 1973). According to this theory, firms convey signals under conditions of information asymmetry to reduce uncertainty and influence stakeholders' perceptions (Williams & Hodges, 2022).

In sustainable consumption research, when consumers perceive sustainability signals as clear and verifiable, they are more likely to regard the brand as credible and reliable (Fu et al., 2023; Sansome et al., 2024). Several studies have implicitly or explicitly applied this perspective by examining how sustainability labels, certifications, brand reputation, and communication cues shape consumer evaluations and purchasing decisions (Rese & Baier, 2025). For instance, sustainability labels and certifications have been shown to function as signals of environmental and ethical quality, helping consumers infer product attributes that cannot be directly assessed. Behre & Cauberghe (2025) demonstrate that third-party certification labels can work as credible signals in online fashion because they reassure consumers that sustainability claims are not merely self-declared by the brand. Similarly, Negash & Akhbar (2024) apply signalling theory to second-hand fashion, arguing that information such as seller reputation, product history, condition details, and refurbishment descriptions can reduce uncertainty and increase trust. Huh & Kim (2022) show that green apparel can operate as a social signal, especially among younger consumers, because wearing sustainable products may communicate status, moral awareness, or belonging to a socially responsible lifestyle.

Therefore, signalling theory is employed in sustainable fashion research in two complementary ways: first, to explain how consumers use labels, certifications, seller information, and product transparency to assess whether sustainability claims are trustworthy; and second, to show how sustainable fashion consumption allows individuals to express identity, values, distinction, and social positioning through what they wear.

3 Method

This study aims to investigate how digital ESG communication shapes consumer responses in the sustainable fashion industry, with particular attention to the role of consumer trust and cross-country differences between

Italy and South Korea. Specifically, the research focuses on how fashion companies use digital channels to communicate their ESG practices, and on how consumers interpret these signals. Accordingly, this study adopts a narrative literature review approach (Baumeister & Leary, 1997). The review followed a structured workflow consisting of three main stages: identification, screening, and content analysis.

During the identification phase, an initial pool of articles was collected through a broad search in Scopus using combinations of keywords, such as “sustainable fashion”, “consumer behaviour”, “ESG communication”, and “ESG practice”.

In the second stage, the abstracts were iteratively screened based on conceptual relevance. Studies were retained when they provided meaningful contributions to at least one of the following domains:

- sustainable fashion consumption,
- ESG communication practices,
- and cross-country evidence for Italy and South Korea.

The screening process was then validated through triangulation among the authors and returned 30 available articles for content analysis.

In the final stage, the final pull of studies was analysed through the combined lenses of signalling theory. This interpretive approach enabled the integration of findings into a coherent model.

4 Results

4.1 ESG communication practices

ESG communication, or sustainability communication, refers to the set of messages, activities, and channels through which a company transparently and consistently conveys its commitment to ESG issues, according to Waites (2025). Through ESG communication, firms inform stakeholders about their environmental actions, their alignment with social and corporate values, and the transparency of their governance and managerial practices. The most recent literature reveals that the increasing relevance of ESG factors has progressively reshaped both corporate strategies and consumer evaluation processes (Waites, 2025) and expectations (Su & Teo, 2025). Specifically, we analysed articles based on their E-S-G focus to distinguish communication related to each pillar and discover their impacts on consumer behaviour.

The environmental dimension emerges as particularly salient in shaping consumer responses. Studies consistently show that consumers tend to favour firms perceived as environmentally responsible, and in some cases are willing to pay a premium for products associated with sustainability, especially when such efforts are clearly communicated (Su & Teo, 2025). However, the evidence also suggests that the effectiveness of environmental claims depends strongly on how information is presented and perceived. Transparent and visible communication enhances trust and ethical satisfaction, thereby strengthening green purchase intentions, whereas low salience or ambiguous communication may limit behavioural conversion despite positive attitudes (Gherghina, 2024).

The social pillar further reinforces the relevance of ESG communication for consumer behaviour. The reviewed studies show that communication related to social responsibility, such as fair labour practices, community engagement, and inclusiveness, can directly influence conscious consumption, particularly among socially aware consumers. At the same time, consumers appear highly sensitive to perceived inconsistencies or violations of social expectations, which may undermine trust and negatively affect behavioural responses (Leonelli et al., 2025). These findings suggest that social ESG communication contributes to consumer trust not only by evidencing positive practices but also by reducing perceived ethical risk.

The governance mechanisms, although less directly visible to consumers, play a key role in reinforcing the credibility of environmental and social claims (Su & Teo, 2025). In this sense, governance acts as a foundational mechanism that supports the perceived reliability of ESG communication, thereby strengthening consumer trust even when governance itself is not explicitly evaluated at the point of purchase. Importantly, the reviewed studies converge in highlighting that the effectiveness of ESG communication depends on the integration of the three dimensions and on the coherence of the overall message. Firms that align ESG practices transparently and consistently are more likely to generate consumer trust, loyalty, and willingness to support the brand (Su & Teo, 2025), thereby amplifying their impact on purchase intention and willingness to pay. Finally, the analysis suggests that consumers who actively engage with ESG-related information tend to exhibit stronger behavioural responses, including loyalty and advocacy behaviours (Leonelli et al., 2025). Figure 1 resumes the main evidence in consumer behaviour literature analysing ESG factors, derived from Tan et al.’s (2025) framework:

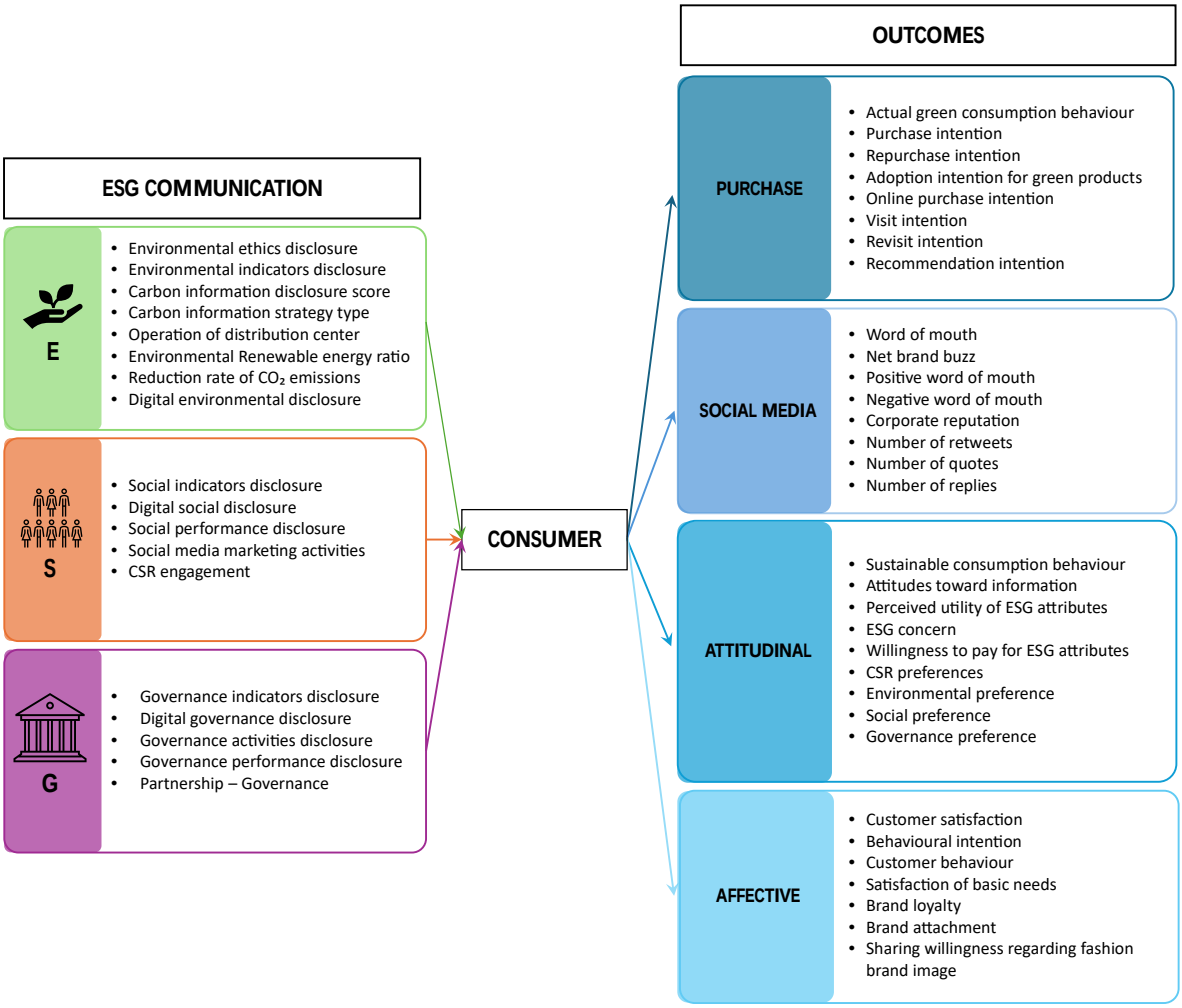


Figure 1: Linking ESG communication with consumer outcomes. Adaptation from Tan et al. (2025).

4.2 Cross-country comparison: Italy and South Korea

Recent ESG-consumer literature shows that consumer responses to sustainability are often context-specific and shaped by differences in regulatory environments, market structures, public awareness, and culturally embedded interpretations of trust, responsibility, and corporate legitimacy. In this respect, Tan et al. (2025) highlight both a marked concentration of studies in East Asia and the need to broaden empirical engagement to strengthen the cross-cultural applicability of existing models. Against this background, focusing on Italy and South Korea is theoretically meaningful. Italy represents a mature fashion context, deeply rooted in manufacturing tradition and closely associated with the cultural legacy of slow fashion, craftsmanship, product quality, and symbolic value. South Korea, by contrast, offers a highly dynamic, digitalised, and competitive market environment in which consumer-brand interactions are more strongly mediated by platforms, technological adoption, and fast-evolving communication practices. To compare the two focal contexts more systematically, the selected studies were grouped into categories based on the variables examined and the main findings reported. A signalling-based interpretation was then added to clarify how each contribution helps explain the reception, decoding, or behavioural effectiveness of sustainability-related cues. Table 1 represents the main findings of the coding process:

Table 1: Extraction of literature coding for cross-country comparison and signalling interpretation.

Category	Country	Key variables	Main finding	Signalling interpretation
Sustainability predisposition	Italy	Environmental knowledge, environmental concern, eco-friendly behaviour	Green purchase intention is positively associated with environmental knowledge, environmental concern, and eco-friendly behaviour (Annunziata et al., 2023).	Consumer predisposition reduces ambiguity in the decoding of sustainability signals.
	South Korea	Idealism, moral intensity	Idealism increases moral intensity, which in turn increases sustainable product purchase behaviour (Hong & Kang, 2019).	Consumer predisposition increases the likelihood that sustainability information is interpreted as meaningful rather than ignored.
Behavioural barriers	Italy	Shopping habits, price sensitivity, second-hand willingness, bio-based preference	Shopping habits shape sustainable fashion uptake; second-hand adoption is weaker than bio-based acceptance (Colasante et al., 2025). Higher prices may lead to greater scepticism towards sustainable products (Fani et al., 2023).	Positive sustainability signals are insufficient when habitual consumption patterns, price orientation, or preference for conventional options dominate the decision process.
	South Korea	Utilitarian value, subjective norms, environmental concern, perceived risk	Intention and behaviour diverge; perceived risk and product-evaluation issues matter for actual uptake Park & Lin (2020).	Behavioural conversion is weakened by perceived risk and product-related uncertainty.
Digital and social influence	Italy	Visible social cues, environmental concern, self-image	Sustainability cues do not always increase preference; visible signalling may reduce preference depending on self-presentation dynamics (Fani et al., 2023).	The effectiveness of sustainability cues depends on whether they are socially acceptable and consistent with the consumer's concerns.
	Italy	Online and traditional purchasing channels	The digital environment is treated mainly as an informational and purchasing channel rather than as a socially embedded relational space (Annunziata et al., 2023).	The digital environment operates primarily as a transmission channel rather than as a relational mechanism reinforcing the signal.
	South Korea	Online interaction, social capital	Social capital strongly affects sustainable fashion purchase intention in a YouTube setting (Kim et al., 2020).	Digital platforms function as socially embedded environments that reinforce signals through interaction and relational validation.
	South Korea, China, Russia	Digital advertising cue, country interpretation	The same digital cue affects consumer evaluation differently across contexts (Whang et al., 2015).	The same signal can be decoded differently, affecting whether it is perceived as credible or persuasive.

Both contexts display a meaningful sustainability predisposition. In Italy, Annunziata et al. (2023) show that environmental knowledge, environmental concern, and eco-friendly behaviour positively affect green purchase intention in the fashion industry. In South Korea, Hong & Kang (2019) show that sustainable fashion consumption is also supported by ethical-cognitive mechanisms, as idealism positively affects moral intensity, which in turn influences purchase behaviour. This suggests that consumers in both countries can be receptive to sustainability information, even though such a predisposition alone does not guarantee behavioural conversion. However, behavioural barriers persist and may explain the well-known attitude-behaviour gap. In Italy, the literature points primarily to economic and habitual constraints. Colasante et al. (2025) show that shopping habits and price orientation remain important barriers to the adoption of more sustainable fashion practices, and related

Italian evidence indicates weaker second-hand adoption and a higher need for economic justification in circular fashion choices. These findings suggest that, even when trust in sustainability communication is present, behavioural response may remain fragile because it competes with price sensitivity, routine purchasing practices, and the enduring attractiveness of fast fashion. In South Korea, by contrast, the barriers appear more perceptual and related to credibility. Park & Lin (2020) and product-evaluation concerns help explain the gap between intention and actual uptake in recycled and upcycled fashion products. Hong & Kang (2019) further indicate that moral intensity matters for purchase behaviour. Hence, while both contexts display barriers to behavioural conversion, these appear more habit-based in Italy and more risk- and perception-based in South Korea. Finally, sustainability claims in fashion retain the nature of credence attributes and therefore require interpretive mechanisms through which information becomes believable. In Italy, online visible sustainability cues do not always strengthen preference and may even backfire when they interfere with self-presentation dynamics, suggesting that trust in sustainability information is more conditional and context-sensitive (Fani et al., 2023). In addition, the digital sphere appears to function mainly as an informational and purchasing environment: Annunziata et al. (2023) distinguish between online and traditional channels when examining green purchase intention, treating digitalisation primarily as a channel condition rather than as a relational space. In South Korea, online cues appear more socially reinforced. Kim et al. (2020) show that, in digital environments, parasocial interaction and especially social capital significantly influence purchase intention for sustainable fashion products, while Whang et al., (2015) show that the same market cue is interpreted differently across countries, indicating that credibility is context-dependent. Thus, in South Korea, digital communication is more embedded in interaction and relational validation than in Italy.

4.3 Empirical model proposal

Building on the previous insights, the proposed empirical model is grounded in signalling theory. This perspective is particularly appropriate in sustainable fashion contexts, where consumers often face information asymmetry and cannot directly assess whether brands' sustainability claims reflect actual ESG practices. However, in digital environments, the mere availability of ESG information is not sufficient to generate favourable consumer responses (Sansome et al., 2024). Accordingly, the model assumes that digital ESG signals influence consumer outcomes through a process of signal interpretation, in line with the literature (Tan et al., 2025). The overall objective is to provide a basis for a future multi-group analysis comparing Italy and South Korea, where the proposed relationships may operate with different intensity across national contexts.

Specifically, ESG signals refer to the set of sustainability-related information and communication practices through which fashion brands disclose their ESG commitments. In line with recent signalling-based studies, these signals may be formal (Behre & Cauberghe, 2025) or informal and market-based (Negash & Akhbar, 2024). Hence, these may include disclosure related to environmental claims, supply-chain information, ethical labour practices, circularity initiatives, certifications, sustainability reports, product-level cues, social media content and messages. ESG signals are therefore not limited to formal reporting but include the broader ecosystem of digital, symbolic, and relational communication through which consumers encounter, evaluate, and make sense of sustainability claims (Fani et al., 2023; Sansome et al., 2024).

As shown in Figure 2, the authors advanced different hypotheses to build the model. First, the positive antecedents of ESG signal interpretation (IS) are derived from the cross-country literature analysis and are conceptualised as two exogenous constructs: sustainability predisposition (SP) and digital and social influence (DSI).

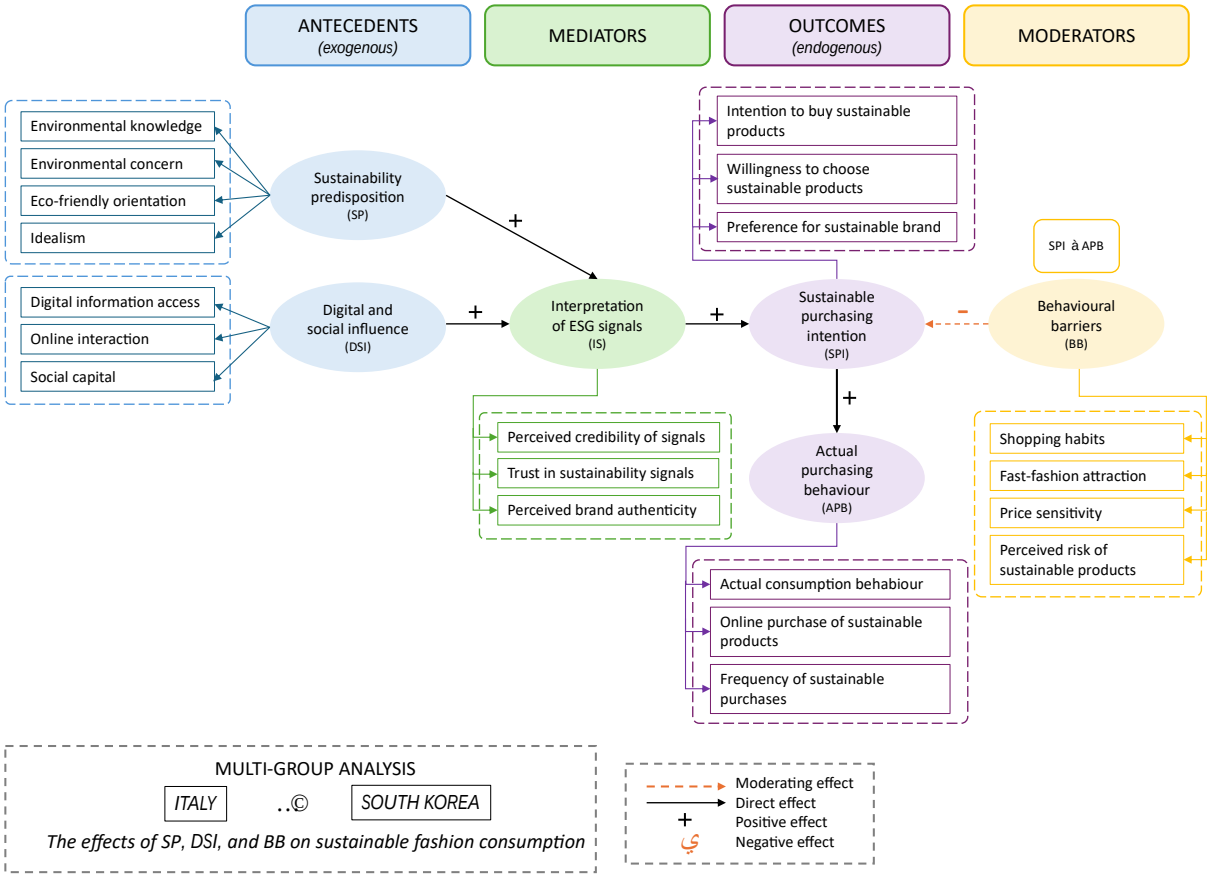


Figure 2: Multi-group empirical model proposal.

SP captures consumers' prior orientation toward sustainability and can be measured through items related to environmental knowledge, environmental concern, eco-friendly behaviour, and idealism. DSI captures the extent to which consumers are exposed to, search for, and validate ESG information through digital and relational channels. This can be measured through items related to digital information access, online interaction, and social capital. These antecedents have a positive influence on the central mediator of the model, which is the interpretation of ESG signals (IS). It captures how consumers make sense of the ESG information provided by fashion brands (Sansome et al., 2024; Tan et al., 2025). It can be measured through items reflecting perceived credibility, trust, and perceived brand authenticity (Fu et al., 2023; Tan et al., 2025). The measurement items derive from theory and literature, which suggest that consumers respond more favourably when sustainability information is perceived as credible, reliable, and aligned with corporate behaviour (Negash & Akhbar, 2024).

Downstream, the model links sustainable purchasing intention (SPI) with actual purchasing behaviour (APB), a positive relation theoretically grounded in the Theory of Planned Behaviour (Ajzen, 1991), according to which intention represents the motivational antecedent of behaviour but does not necessarily guarantee actual action.

Finally, the model introduces behavioural barriers (BB) as negative moderators of the relationship between SPI and APB. These barriers emerged from the cross-country analysis and explain why consumers may not translate favourable intentions into actual sustainable purchases. They include price sensitivity, fast-fashion habits, and perceived risk (as hygiene concerns). Therefore, the model assumes that even when ESG signals are positively interpreted and generate sustainable purchasing intention, BB may reduce the likelihood that such intention becomes actual purchasing behaviour. Thus, the model does not assume that intention automatically translates into behaviour; rather, it explicitly accounts for the intention-behaviour gap that often characterises sustainable fashion consumption.

In this sense, the model integrates ESG communication, consumer interpretation, and behavioural constraints into a single framework suitable for future cross-country empirical testing.

5 Conclusion

This paper addresses the gap concerning how digital ESG communication practices influence sustainable fashion consumption. To address this issue, the study adopts a narrative literature review aimed at integrating

research on ESG communication practices, sustainable fashion consumption, and cross-country differences, focusing on the comparison between Italy and South Korea.

The review shows that ESG communication may not automatically generate trust or sustainable purchasing behaviour. Rather, its effectiveness depends on consumers' ability to interpret ESG information as credible, authentic, and meaningful. Building on signalling theory, the authors therefore develop an empirical model in which digital ESG communication is conceptualised as a set of sustainability signals whose behavioural impact depends on consumers' interpretation of such signals. This interpretation of ESG signals represents the central construct of the model, as it explains how ESG communication becomes relevant for consumer decision-making.

The cross-country analysis further allows the identification of three key constructs: sustainability predisposition (SP), digital and social influence (DSI), and behavioural barriers (BB). SP and DSI are proposed as antecedents of signal interpretation, while BB moderates the relationship between sustainable purchasing intention and actual purchasing behaviour. In conclusion, the model provides a theoretical basis for future empirical testing through multi-group analysis and helps explain how ESG communication may operate differently across national contexts.

5.1 Implications of the study

From a theoretical perspective, the paper extends the application of signalling theory in sustainable fashion consumption by specifying digital ESG communication as a distinct signalling environment. Second, the study advances a more context-based view of sustainable fashion consumption by showing that the same model cannot be assumed to operate identically across national settings. The cross-country comparison further refines this contribution by identifying three contextual dimensions through which Italy and South Korea differ regarding sustainability predisposition, digital and social influence, and behavioural barriers. These differences support the idea that the proposed model retains a common structure across countries, while its internal paths are expected to operate with different intensity across contexts.

The paper also offers relevant practical implications for managers and practitioners in the fashion industry. For fashion firms, the findings suggest that digital ESG communication should not simply increase the quantity of sustainability disclosures but should improve their interpretability, transparency, and credibility. In practical terms, this means moving beyond vague claims and symbolic sustainability rhetoric toward clearer and more verifiable communication formats, such as traceability tools, standardised indicators, and coherent product-level disclosures. However, the model also suggests that the effectiveness of these tools is likely to differ across contexts. In Italy, brands may need to complement ESG communication with clearer demonstrations of consumer value, convenience, and practical relevance, because transparency alone may not overcome economic and habitual barriers. In South Korea, by contrast, digital ESG communication may be more effective when embedded in interactive and socially mediated environments and when supported by visible and credible verification cues.

5.2 Limits and future research directions

The study presents some important limitations. First, the paper is conceptual and does not provide empirical testing of the proposed relationships. Second, the content analysis and the following cross-country comparison rely on a focused and purposive interpretation of existing studies. Third, the study concentrates on the fashion industry; therefore, the generalisability of the model to other ESG-intensive industries should not be assumed automatically. These limitations open several directions for future research. A first direction concerns the operationalisation of digital ESG communication itself, which future studies could unpack into more specific communication formats. A second avenue concerns the empirical validation of the model through multi-group comparative designs between Italy and South Korea. Finally, future research could broaden the contextual scope beyond Italy and South Korea to respond to recent calls for wider contextual coverage in ESG consumer research and to test whether the proposed model holds across other cultural and institutional settings.

6 Ethics declaration

No ethical clearance was required for this study.

7 AI declaration

The authors declare the use of ChatGPT to support language refinement.

References

Annunziata, E., Pucci, T., Cammeo, J., Zanni, L., & Frey, M. (2023). The mediating role of exogenous shocks in green purchase intention: evidence from italian fashion industry in the Covid-19 era. *Italian Journal of Marketing*, *2023*(1), 59–79. https://doi.org/10.1007/s43039-023-00065-4

Baumeister, R. F., & Leary, M. R. (1997). Writing Narrative Literature Reviews. In *Review of General Psychology* (Vol. 1, Number 3).

Behre, B., & Cauberghe, V. (2025). Signaling sustainability in online fashion consumption: The role of credibility induced by certification labels. *Cleaner and Responsible Consumption*, *17*. https://doi.org/10.1016/j.clrc.2025.100287

Colasante, A., D'Adamo, I., Rosa, P., & Morone, P. (2025). How consumer shopping habits affect willingness to embrace sustainable fashion. *Applied Economics Letters*, *32*(6), 850–855. https://doi.org/10.1080/13504851.2023.2290578

Fani, V., Mazzoli\, V., & Acuti, D. (2023). *"I wannabe sustainable, but I don't wanna show it!": The effect of sustainability cues on young adult consumers' preferences*.

Fu, S., Ma, R., He, G., Chen, Z., & Liu, H. (2023). A study on the influence of product environmental information transparency on online consumers' purchasing behavior of green agricultural products. *Frontiers in Psychology*, *14*. https://doi.org/10.3389/fpsyg.2023.1168214

Gherghina, Ştefan C. (2024). Corporate Finance and Environmental, Social, and Governance (ESG) Practices. In *Journal of Risk and Financial Management* (Vol. 17, Number 7). Multidisciplinary Digital Publishing Institute (MDPI). https://doi.org/10.3390/jrfm17070308

Hong, H., & Kang, J. H. (2019). The impact of moral philosophy and moral intensity on purchase behavior toward sustainable textile and apparel products. *Fashion and Textiles*, *6*(1). https://doi.org/10.1186/s40691-019-0170-8

Hörisch, J., Schaltegger, S., & Freeman, R. E. (2020). Integrating stakeholder theory and sustainability accounting: A conceptual synthesis. *Journal of Cleaner Production*, *275*. https://doi.org/10.1016/j.jclepro.2020.124097

Huh, J., & Kim, N. L. (2022). *Green as the New Status Symbol: Examining Green Signaling Effects among Gen Z and Millennial Consumers*.

Kim, J., Kang, S., & Lee, K. H. (2020). How social capital impacts the purchase intention of sustainable fashion products. *Journal of Business Research*, *117*, 596–603. https://doi.org/10.1016/j.jbusres.2018.10.010

Koszewska, M., Rahman, O., & Dyczewski, B. (2020). Circular Fashion - Consumers' Attitudes in Cross-National Study: Poland and Canada. *Autex Research Journal*, *20*(3), 327–337. https://doi.org/10.2478/aut-2020-0029

Leonelli, S., Muhn, M., Rauter, T., & Sran, G. S. (2025). How do consumers use ESG disclosure? Evidence from a randomized field experiment with everyday product purchases. *Journal of Accounting and Economics*. https://doi.org/10.1016/j.jacceco.2025.101811

Market, J., Author, S., & Spence, M. (1973). Job Market Signaling. In *Source: The Quarterly Journal of Economics* (Vol. 87, Number 3).

Negash, Y. T., & Akhbar, T. (2024). Building consumer trust in secondhand fashion: A signaling theory perspective on how consumer orientation and environmental awareness shape engagement. *Cleaner and Responsible Consumption*, *14*. https://doi.org/10.1016/j.clrc.2024.100211

Park, H. J., & Lin, L. M. (2020). Exploring attitude–behavior gap in sustainable consumption: comparison of recycled and upcycled fashion products. *Journal of Business Research*, *117*, 623–628. https://doi.org/10.1016/j.jbusres.2018.08.025

Rese, A., & Baier, D. (2025). Rental clothing box subscription: The importance of sustainable fashion labels. *Journal of Retailing and Consumer Services*, *83*. https://doi.org/10.1016/j.jretconser.2024.104153

Ronda, L. (2024). Overcoming barriers for sustainable fashion: bridging attitude-behaviour gap in retail. *International Journal of Retail and Distribution Management*, *52*(1), 44–61. https://doi.org/10.1108/IJRDM-02-2023-0056

Sansome, K., Wilkie, D., & Conduit, J. (2024). Beyond information availability: Specifying the dimensions of consumer perceived brand transparency. *Journal of Business Research*, *170*. https://doi.org/10.1016/j.jbusres.2023.114358

Su, T., & Teo, P.-C. (2025). A Literature Review of Understanding Consumer Behavior in ESG through Stimulus-Response (S-R) Theory. *International Journal of Academic Research in Business and Social Sciences*, *15*(4). https://doi.org/10.6007/ijarbss/v15-i4/25263

Tan, Q. L., Hashim, S., Abdullah, N. L., Zaki, H. O., & Zheng, Z. (2025). Environmental Social Governance (ESG) and Consumer Behavior: A Systematic Review and Research Agenda. In *International Journal of Consumer Studies* (Vol. 49, Number 5). John Wiley and Sons Inc. https://doi.org/10.1111/ijcs.70105

Tripopsakul, S., & Puriwat, W. (2022). Understanding the Impact of ESG on Brand Trust and Customer Engagement. *Journal of Human, Earth, and Future*, *3*(4). www.HEFJournal.org

Waites, S. F. (2025). Decoding ESG: Consumer Perceptions, Ethical Signals and Financial Outcomes. *Journal of Risk and Financial Management*, *18*(7). https://doi.org/10.3390/jrfm18070361

Whang, H., Ko, E., Zhang, T., & Mattila, P. (2015). Brand popularity as an advertising cue affecting consumer evaluation on sustainable brands: A comparison study of Korea, China, and Russia. *International Journal of Advertising*, *34*(5), 789–811. https://doi.org/10.1080/02650487.2015.1057381

Williams, A., & Hodges, N. (2022). Signaling Sustainability: Exploring Consumer Perspectives on Communicating Apparel Sustainability Information. *Journal of Sustainable Marketing*, *3*(1), 26–39. https://doi.org/10.51300/jsm-2022-49

TRACING: A Design-Oriented Framework for Enabling Governance Mechanisms in Cold Chain Networks

Danilo Martino[1], Giulia Caso[2], Guglielmo Elefante[1], Claudio Pascarelli[2], Gabriele Papadia[3], Mariangela Lazoi[2]
[1]2038 Innovation Company Srl, Palermo, Italy
[2]University of Salento, Lecce, Italy
[3]Advantech Srl, Lecce, Italy

Abstract

The management of Cold Chain (CC) logistics is critical to ensuring the safety, quality, and integrity of perishable food products, as even minor deviations from controlled conditions can cause irreversible deterioration and economic losses. However, current monitoring practices, often manual or semi-automated, show limited capability in handling real-time data, constraining risk mitigation and regulatory compliance. The increasing adoption of Internet of Things (IoT) technologies and advanced analytics provides new opportunities to enhance visibility and control across distributed supply chains.

This paper presents TRACING (innovaTive AI fRAmework for Cold chaIN Governance), a design-oriented framework that supports governance in Cold Chain networks through the integration of Data and Process Governance. Following a design science research approach, TRACING is conceptualized as a socio-technical artifact enabling decision-making, coordination, and accountability among multiple stakeholders.

The framework comprises four components: Data Governance (data collection, validation, and quality assessment), Process Governance (process modeling, monitoring, and optimization), Data Analytics (leveraging Process Mining and Artificial Intelligence), and an event-driven architecture enabling real-time communication and interoperability. These components enable the transformation of raw sensor data into structured knowledge, supporting continuous performance monitoring and anomaly detection.

The framework is illustrated through a Cold Chain scenario involving controlled product maturation, showing how the integration of data and process perspectives enables transparency, traceability, auditability, and coordination, ultimately enhancing operational reliability and resilience.

This work bridges the gap between data infrastructures and governance requirements, providing a structured approach to support knowledge-driven decision-making in complex supply chain environments.

Keywords: Cold Chain; Data Governance; Process Governance; IoT; Food Supply Chain

Paper type: Academic Research Paper

1 Introduction

The management of Cold Chain (CC) logistics is a critical challenge in perishable food supply chains, where maintaining controlled environmental conditions is essential to ensure product quality, safety, and regulatory compliance. Even minor deviations can cause irreversible deterioration and significant economic losses (Aung and Chang, 2014), a risk amplified by the increasing complexity of globalized supply chains involving multiple distributed actors.

In this context, the Cold Chain can be viewed as a complex socio-technical system characterized by strong interdependencies, strict environmental constraints, and the need for real-time decision-making. Temperature control and traceability are key to ensuring safety and quality (Aung and Chang, 2014), while inefficiencies in monitoring and management contribute significantly to food loss and waste (FAO, 2019).

Despite existing regulatory frameworks, current monitoring practices often rely on manual or semi-automated procedures based on periodic data collection and fragmented systems. This limits real-time anomaly detection, reduces process visibility, and constrains timely and effective decision-making, which remains largely reactive.

The diffusion of digital technologies, such as Internet of Things (IoT), Artificial Intelligence (AI), and advanced analytics, offers new opportunities to improve Cold Chain monitoring and management (Atzori et al., 2010; Ben-Daya et al., 2019). IoT enables continuous data collection, while analytical methods support pattern and anomaly detection. At the same time, Business Process Management (BPM) provides structured approaches for process modeling and optimization (Dumas et al., 2013). Their integration enhances efficiency and resource utilization (Bocciarelli et al., 2023; Ornari et al., 2025). In addition, Process Mining techniques support process discovery, conformance checking, and performance analysis through the extraction of knowledge from event logs (Striani

et al., 2022), addressing limitations of traditional analysis and simulation approaches (Oldenburg et al., 2025). However, their integration into unified governance-oriented frameworks remains limited.

Existing studies have mainly focused on either data-centric approaches, such as IoT-based monitoring and traceability systems, or process-oriented approaches based on BPM and logistics optimization (Aung and Chang, 2014; Atzori et al., 2010; Ben-Daya et al., 2019; Jedermann et al., 2014; Verdouw et al., 2016). While these contributions advance specific aspects of Cold Chain operations, they often lack integration between data and process governance, limiting the ability to transform data into actionable knowledge for coordination, accountability, and real-time decision-making (Khatri and Brown, 2010; Otto, 2011).

To address these limitations, this paper proposes TRACING, a design-oriented framework aimed at enabling governance mechanisms in Cold Chain networks through the integration of data and process perspectives. The study adopts a Design Science Research approach (Peffers et al., 2007), complemented by the use of a representative application context to illustrate the design and functionality of the proposed artifact.

TRACING is conceptualized as a socio-technical artifact that structures the collection, management, and analysis of data while enabling the modeling and monitoring of operational processes. By integrating Data Governance components with Process Governance mechanisms and advanced analytics techniques, the framework provides a unified environment for managing complex Cold Chain operations.

2 Methodology

This study adopts a Design Science Research (DSR) approach, which is particularly suitable for addressing complex organizational and technological problems through the design and development of innovative artifacts. Design Science Research focuses on the creation and evaluation of artifacts—such as models, methods, and systems—that contribute both to solving practical problems and to advancing scientific knowledge (Peffers et al., 2007). In the context of this research, the TRACING framework is conceptualized as a socio-technical artifact designed to enable governance mechanisms in Cold Chain networks.

The adoption of a DSR approach is motivated by the nature of the problem addressed in this study. Cold Chain environments are characterized by high operational complexity, distributed responsibilities, heterogeneous data sources, and strict requirements for monitoring, traceability, and regulatory compliance. These characteristics require not only analytical investigation but also the structured design of integrated solutions capable of supporting coordination and decision-making across multiple actors, particularly in contexts where data governance and decision rights play a critical role (Khatri and Brown, 2010; Otto, 2011). Therefore, rather than focusing on hypothesis testing or statistical generalization, this research aims to develop a coherent framework that addresses a class of problems related to governance in data-intensive and process-driven environments.

The methodological process follows a structured sequence of activities aligned with Design Science Research principles. First, the problem context is identified through the analysis of Cold Chain operations, with particular attention to challenges related to environmental monitoring, data fragmentation, and process coordination. Second, the objectives of the artifact are defined in relation to the need for improving visibility, traceability, and accountability within distributed supply chains. Third, the artifact is designed by integrating multiple components, including Data Governance mechanisms, Process Governance approaches, and advanced analytical capabilities. Finally, the framework is illustrated through a representative application context in order to demonstrate its structure and potential applicability.

The study adopts a Design Science Research approach (Peffers et al., 2007), complemented by the use of a representative application context to illustrate the design and functionality of the proposed artifact. The selected context refers to a Cold Chain scenario involving the management of perishable products, where environmental conditions and process execution must be continuously monitored and controlled. Within this perspective, the application context is not used for empirical validation in a strict case study sense, but rather as an illustrative domain that allows demonstrating how the framework can be instantiated in practice.

The design of the TRACING framework is grounded in both theoretical and domain-specific knowledge, including contributions on data governance structures, decision rights, and knowledge management in complex organizational environments (Khatri and Brown, 2010; Otto, 2011; Nonaka, 1994; Alavi and Leidner, 2001). From a theoretical perspective, the framework draws on Business Process Management principles for the modeling and monitoring of workflows (Dumas et al., 2013), as well as on process-oriented analytical approaches such as Process Mining. From a domain perspective, the design is informed by knowledge related to Cold Chain operations, including requirements for temperature control, traceability, and compliance with food safety regulations.

The evaluation of the framework is conducted at a conceptual level, focusing on its capability to enable governance mechanisms such as transparency, traceability, accountability, and coordination. This evaluation is based on the internal coherence of the artifact, on the alignment between identified problems and proposed solutions, and on the logical integration of its components. While no quantitative performance assessment is provided, the framework is assessed in terms of its potential to support decision-making and improve process visibility in complex environments.

By adopting a design-oriented methodology supported by an illustrative application context, this study ensures both rigor and relevance. The approach allows developing a structured and theoretically grounded framework while maintaining a strong connection to practical challenges in Cold Chain management.

3 Case Description: Cold Chain Process

The application context considered in this study refers to a Cold Chain scenario involving the management and monitoring of perishable food products, where maintaining controlled environmental conditions throughout storage, transportation, and distribution is essential. Cold Chain logistics encompasses a sequence of interconnected activities aimed at preserving product quality and safety, particularly for temperature-sensitive goods, whose deterioration may occur rapidly in the presence of environmental deviations (Aung and Chang, 2014; Jedermann et al., 2014; Ruiz-Garcia et al., 2010). The process is characterized by the involvement of multiple actors operating across different stages of the supply chain, including producers, logistics operators, storage facilities, and distribution entities. Each actor contributes to the execution of specific activities, which collectively determine the effectiveness of the overall system. The distributed nature of these operations introduces significant coordination challenges, particularly in relation to the synchronization of activities and the management of information flows across organizational boundaries.

The Cold Chain process can be represented through a controlled maturation workflow, which highlights the sequence of operational phases and the interactions among the involved actors. The process includes key stages such as product reception, storage under controlled conditions, maturation management, and distribution. Each phase is associated with specific environmental requirements and operational constraints, requiring continuous monitoring and precise coordination to ensure product integrity.

For this reason, the technology platform must be able to provide alerts useful for monitoring critical temperature values and parameters, enabling real-time monitoring and supporting any corrective actions. Similarly, the platform must support analysis of historical data and critical parameters useful for effective management of the ripening process. Currently, the Wholesaler/Ripener, while retaining this data, does not perform this type of analysis, which could serve as a basis for any improvements compared to previous practices.

Furthermore, the platform will allow for the clarification and formalization of procedures that currently rely heavily on the experience and tacit knowledge of expert operators, who have consolidated the various approaches currently adopted over the years. Finally, the platform will enable a clear and formalized end-to-end view of the ripening process, which can sometimes be a black box for the operators involved.

To further clarify the operational context, the Cold Chain scenario can be described in terms of actors, data generation, and decision-making processes. The main actors involved include producers, logistics operators, storage facilities, and distribution entities, each responsible for specific stages of product handling and environmental control.

Data are generated from both sensor-based systems and operational activities. Sensor data primarily include environmental variables such as temperature, humidity, and storage conditions, continuously monitored throughout the lifecycle of the product. In parallel, event data are produced by process execution, including product reception, storage transitions, handling operations, and shipment events.

These data streams are integrated within the platform to enable the monitoring of key process indicators, such as temperature compliance, duration of storage phases, occurrence of deviations, and frequency of rework or corrective actions. Based on these inputs, the system supports decision-making processes such as anomaly detection, alert generation, process adjustment, and coordination among actors.

Responsibilities are distributed across actors according to process roles, with each stakeholder accountable for specific operational phases and data inputs. This structured integration of data and responsibilities enables traceability and supports compliance with regulatory requirements, while providing a concrete instantiation of the TRACING framework within a Cold Chain context.

A defining feature of Cold Chain processes is the critical role played by environmental variables, particularly temperature, which must be continuously monitored and maintained within predefined thresholds. Deviations from these thresholds can lead to irreversible damage, affecting product quality and safety, and increasing the risk of food loss and waste (FAO, 2019). In addition to temperature, other factors such as humidity and

atmospheric conditions may influence product preservation, further increasing the complexity of process management.

The operational complexity of the Cold Chain is further amplified by the dynamic and uncertain conditions in which these processes take place. Transportation phases are subject to delays, equipment failures, and environmental fluctuations, which may compromise the stability of storage conditions (Yadav et al,2025). Moreover, the involvement of multiple actors and heterogeneous systems leads to data fragmentation, as information is collected through different devices and stored in separate repositories. This fragmentation limits process visibility and reduces the ability of stakeholders to detect anomalies and respond in a timely manner.

From a process perspective, Cold Chain operations generate both sensor data and event data. Sensor data derive from monitoring devices that capture environmental conditions, while event data are associated with the execution of operational tasks. The integration of these data sources is essential to enable a comprehensive understanding of process behavior, as it allows linking environmental conditions to specific operational activities. However, in many real-world scenarios, such integration remains limited, creating a gap between data availability and actionable knowledge.

Another critical aspect concerns regulatory compliance. Cold Chain operations are subject to strict requirements related to the monitoring and recording of environmental conditions at critical control points. The absence of continuous and reliable monitoring mechanisms may lead to non-compliance, exposing organizations to regulatory and reputational risks. In this context, the need for integrated solutions that combine data management, process monitoring, and analytical capabilities becomes particularly evident.

The Cold Chain scenario considered in this study is therefore used as a representative application context to illustrate the relevance and applicability of the proposed TRACING framework. Rather than focusing on quantitative performance outcomes, the analysis emphasizes the structural characteristics of the process and the challenges associated with data integration, coordination, and governance in distributed environments.

From a regulatory perspective, Cold Chain operations are subject to strict requirements related to food safety, traceability, and environmental control. International standards and regulatory frameworks, such as HACCP (Hazard Analysis and Critical Control Points), ISO 22000 for food safety management systems, and European regulations on food traceability (e.g., Regulation (EC) No 178/2002), impose obligations regarding the monitoring and recording of critical parameters throughout the supply chain. These frameworks require the identification of critical control points, continuous monitoring of environmental conditions, and the availability of auditable records. In this context, the integration of data and process governance, as enabled by the TRACING framework, directly supports compliance by ensuring data consistency, traceability, and accountability across all stages of the Cold Chain.

4 The TRACING framework

The TRACING framework is designed as a socio-technical artifact aimed at enabling governance mechanisms in Cold Chain environments through the integration of data management, process monitoring, and advanced analytics. The framework addresses the limitations identified in the previous sections by providing a unified architecture that links data flows, process execution, and analytical capabilities within a coherent system.

The TRACING framework is structured into main components: Data Governance, Process Governance, and Data Analytics. These components interact within a modular and scalable environment and are supported by an Event Driven Architecture, which enables interoperability among heterogeneous systems and supports real-time data processing.

The Data Governance component is responsible for the collection, ingestion, validation, and quality assessment of data generated throughout the Cold Chain, ensuring data reliability and supporting decision-making processes through clearly defined governance mechanisms (Khatri and Brown, 2010; Otto, 2011; Weber et al., 2009). Data are acquired from multiple sources, including sensors and operational systems, enabling continuous monitoring of environmental conditions through IoT-based infrastructures (Atzori et al., 2010; Ben-Daya et al., 2019; Verdouw et al., 2016). Data quality mechanisms ensure consistency, completeness, and reliability, enabling the use of trustworthy data for subsequent analysis.

The Process Governance component enables the modeling and monitoring of operational workflows. Based on Business Process Management principles (Dumas et al., 2013), processes are represented as structured sequences of activities, with clearly defined roles and responsibilities. Workflow monitoring tools allow stakeholders to visualize process execution and identify deviations, while rule-based mechanisms support the generation of alerts and corrective actions.

The Data Analytics component integrates advanced analytical techniques aimed at extracting actionable insights from operational data, particularly through the use of Process Mining methods for process discovery,

conformance checking, and performance analysis (van der Aalst, 2016; van der Aalst et al., 2012; Rojas et al., 2016). These techniques enable the reconstruction of actual process flows, the identification of deviations from expected behaviors, and the evaluation of performance indicators such as delays and rework activities.

The Event Driven Architecture enables real-time communication and coordination among system components. Through an event-based model, the framework supports asynchronous data exchange, ensuring scalability and responsiveness. Events generated by sensors or process activities trigger automated responses, such as alerts and notifications, allowing stakeholders to react promptly to changes in operational conditions.

The integration of these components enables the transformation of raw data into structured knowledge, supporting decision-making and governance. By linking data generation, process execution, and analytical interpretation, TRACING provides a comprehensive and dynamic view of Cold Chain operations, overcoming the fragmentation and limitations of traditional monitoring systems.

To provide a more operational and implementation-oriented representation of the framework, the components described above are further articulated into a layered architecture, shown in Figure 1, which details their functionalities and interactions within the technology platform.

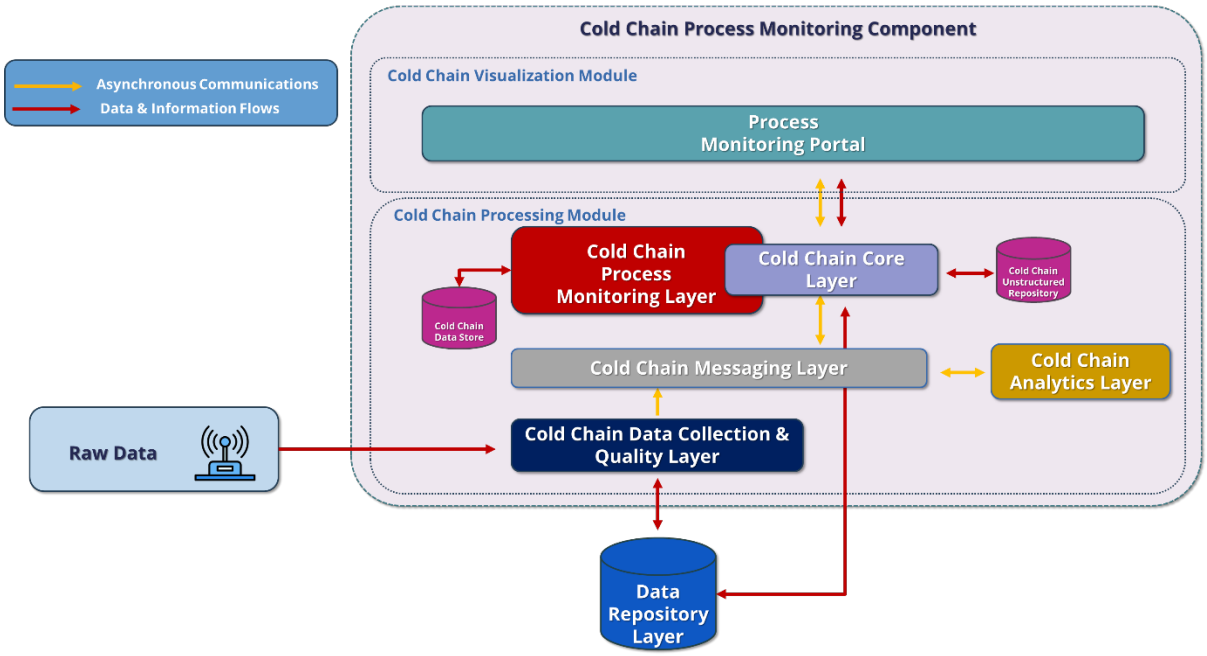


Figure 1: Architecture of the enabling technology platform

5 Results

The platform was applied to a real-world use case at Interfrutta Srl Fatano 1880, a company managing, among others, the maturation and distribution of climacteric fruits. The case demonstrated the capability of TRACING to support controlled maturation cycles by monitoring key variables and providing actionable alerts to prevent spoilage or suboptimal product quality. While the study does not empirically assess governance outcomes, it illustrates how TRACING structures information flows to enable mechanisms such as accountability through data provenance, coordination through process modeling, auditability through event logging, and transparency through shared dashboards. These elements collectively create conditions necessary for transforming raw IoT data into intelligent knowledge usable across networked actors in the Cold Chain.

The process under analysis concerns the ripening operations related to use case. After verifying transport conditions through the analysis of recorded temperature data, the lots of bananas are identified in the company's information systems and sorted according to orders, then addressed either to the ripening phase or to storage. In both cases, they are placed in temperature-controlled rooms: storage is temporary process, while ripening is a process that evolves over time and is managed in a controlled manner throughout all its stages.

During the process, data related to the setpoint temperature values defined during both the ripening and storage phases, the monitored temperature values, and variables related to ethylene emission or air extraction from the chambers are stored within the information systems and converted into log files. Process Mining techniques are applied to these log files.

Process Discovery allows to obtain an initial model on which to analyze performance, as shown in Figure 2. In particular, the focus is on calculating rework for each activity in the ripening process in order to identify potential bottlenecks. Figure 2 reports the number of instances in which activities are reworked more than twice.

ACTIVITY	REWORK
HEATING	9
OPENING RIPINING CHAMBER	16
AIR EXTRACTION	33
ETYLENE EMISSION	8
AIR VALVE CLOSURE	15

Figure 2: Performance analysis highlighting rework activities

Subsequently, process variants that lead to bottlenecks are removed, and the resulting model is the one shown in Figure 3, obtained by reapplying the Discovery algorithm, Inductive Miner (Leemans et al., 2013), to the filtered log.

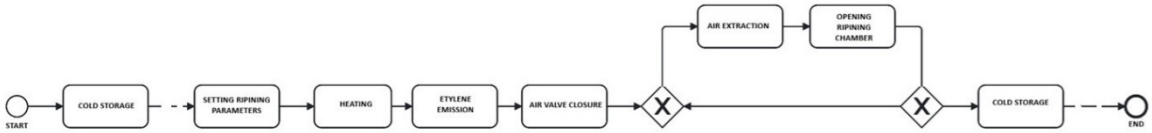


Figure 3: Process Model calculated by Process Discovery algorithm

The validity of the model obtained for the process under consideration is assessed through the application of the Conformance Checking algorithm, specifically Alignment (Van Zelst et al, 2020), which highlights deviations of subsequent acquired log files from the derived process model by computing metrics such as fitness, precision, generalization, and simplicity (Figure 4).

Average Fitness	0.958
Precision	0.561
Generalization	0.778
Simplicity	0.746

Figure 4: Calculated metrics by the Conformance Checking

It can be observed that the model is quite accurate, as the fitness is high, with a moderate level of precision since the model accounts for additional behaviors that are not always present in the real log. Furthermore, it is not rigid, as it shows good generalization, and it is very simple.

6 Enabling Governance Mechanisms

The TRACING framework enables governance mechanisms by structuring the interaction between data, processes, and analytical capabilities. In complex environments such as the Cold Chain, governance emerges from the ability to ensure visibility, coordination, and accountability across distributed actors.

A first key mechanism is transparency, which refers to the ability of stakeholders to access and interpret relevant information regarding process execution and environmental conditions. By integrating data collection and centralized data management, the framework ensures continuous availability of information, reducing uncertainty and enabling informed decision-making.

Traceability is another fundamental mechanism, allowing the tracking of products and processes across the entire supply chain. By linking sensor data with event logs, TRACING enables the reconstruction of process histories, supporting both operational monitoring and retrospective analysis (Aung and Chang, 2014).

Accountability is supported through the explicit representation of roles, responsibilities, and activities within workflows. The alignment between process execution and data records enables the identification of actions performed by specific actors, enhancing responsibility attribution and supporting performance evaluation.

The framework also supports auditability, enabling the reconstruction and verification of past operations. The systematic collection of data and events creates auditable records that facilitate compliance with regulatory requirements and quality standards.

Coordination among actors is enhanced through the event-driven architecture, which supports real-time communication and synchronization of activities. By enabling timely responses to process deviations, the framework improves operational efficiency and reduces risks associated with delays and errors.

This contribution is particularly relevant within the Knowledge Management domain, as it demonstrates how digital infrastructures can support the transformation of data into organizational knowledge through processes of knowledge creation and sharing (Nonaka, 1994; Alavi and Leidner, 2001; Davenport and Prusak, 1998).

By structuring and contextualizing data within process execution, TRACING enables the generation of actionable insights that support decision-making and continuous improvement.

7 Discussion

The TRACING framework contributes to bridging the gap between technological infrastructures and governance-oriented approaches in supply chain management. By integrating data, processes, and analytics within a unified system, the framework enables a more comprehensive understanding of complex operations and supports the emergence of governance mechanisms.

From a theoretical perspective, the framework extends existing research on IoT and Business Process Management by explicitly linking data flows and process execution. This integration enables the transformation of data into knowledge, supporting decision-making and organizational learning.

From a practical perspective, TRACING provides a structured approach for designing digital solutions in Cold Chain environments. The framework enhances visibility, supports real-time decision-making, and improves coordination among actors.

8 Conclusions and Future Research

This paper presented TRACING (innovaTive AI fRAmework for Cold chaIN Governance), a design-oriented framework aimed at enabling governance mechanisms in Cold Chain networks through the integration of data management, process monitoring, and advanced analytics. Grounded in a Design Science Research approach, the study conceptualizes TRACING as a socio-technical artifact capable of addressing the complexity of Cold Chain environments, which are characterized by distributed actors, strict environmental constraints, and the need for real-time decision-making.

The proposed framework contributes to the literature by bridging the gap between data-centric technological infrastructures and governance-oriented approaches in supply chain management. By integrating Data Governance, Process Governance, and analytical components within a unified architecture, TRACING enables key governance mechanisms, including transparency, traceability, accountability, auditability, and coordination. These mechanisms are essential for improving visibility, supporting compliance, and enhancing decision-making in complex and dynamic operational contexts.

A key contribution of this work lies in the structured integration of heterogeneous components, which allows transforming raw data into actionable knowledge. Through the alignment of data flows and process execution, the framework supports a more comprehensive understanding of Cold Chain operations and facilitates the identification of inefficiencies and risks. In this sense, TRACING provides a conceptual foundation for the development of digital solutions that go beyond simple monitoring, enabling knowledge-driven governance in distributed supply chain environments.

From a practical perspective, the framework offers guidance for the design and implementation of integrated platforms capable of supporting real-time monitoring, coordination among actors, and informed decision-making. The modular and scalable nature of the architecture makes it adaptable to different organizational settings and potentially applicable beyond the specific Cold Chain scenario considered in this study.

While the framework is designed to be generalizable, its applicability to other domains requires further investigation.

Future research should focus on exploring the extension of TRACING to other types of supply chains or industrial contexts, as well as the integration of advanced analytical techniques, such as predictive and prescriptive models, to enhance its decision-support capabilities. Finally, the role of organizational and institutional factors, including data governance policies, user adoption, and inter-organizational collaboration, should be investigated to better understand the conditions required for successful implementation.

9 Acknowledgements

The research was part of the research project titled “TRACING - innovaTive AI fRAmework for Cold chaIN Governance”, financed by the cascading call for proposals for research activities conducted by companies,

including in collaboration with public universities, legally recognized private universities accredited by the Ministry of Education, Universities and Research (MUR), and public research institutions, within the scope of Spoke 3 "FOOD SAFETY OF TRADITIONAL AND NOVEL FOODS" of the project "ON Foods (Research and Innovation Network on Food and Nutrition Sustainability, Safety, and Security – Working ON Foods)".

10 Ethics Declaration

This research does not involve human participants, personal data, or any form of experimentation requiring ethical approval. Therefore, ethical clearance was not required for the purposes of this study.

11 AI Declaration

AI-based tools were used for the sole purpose of paraphrasing and language refinement of this manuscript. All content has been manually curated and thoroughly checked for validity, accuracy, appropriateness. All conceptual content, methodological design, references, and scientific contributions are the result of the authors' original work.

References

Alavi, M. and Leidner, D.E. (2001) 'Review: Knowledge management and knowledge management systems: Conceptual foundations and research issues', *MIS Quarterly*, 25(1), pp. 107–136.

Atzori, L., Iera, A. and Morabito, G. (2010) 'The Internet of Things: A survey', *Computer Networks*, 54(15), pp. 2787–2805.

Aung, M.M. and Chang, Y.S. (2014) 'Traceability in a food supply chain: Safety and quality perspectives', *Food Control*, 39, pp. 172–184.

Ben-Daya, M., Hassini, E. and Bahroun, Z. (2019) 'Internet of things and supply chain management: a literature review', *International Journal of Production Research*, 57(15–16), pp. 4719–4742.

Bocciarelli, P., D'ambrogio, A., Panetti, T., (2023) 'A Model Based Framework for IoT-Aware Business Process Management'. Future Internet, 15, p.50.

Davenport, T.H. and Prusak, L. (1998) *Working Knowledge: How Organizations Manage What They Know*. Boston: Harvard Business School Press.

Dumas, M., La Rosa, M., Mendling, J. and Reijers, H.A. (2013) *Fundamentals of Business Process Management*. Berlin: Springer.

FAO (2019) *The State of Food and Agriculture 2019: Moving forward on food loss and waste reduction*. Rome: Food and Agriculture Organization.

Jedermann, R., Nicometo, M., Uysal, I. and Lang, W. (2014) 'Reducing food losses by intelligent food logistics', *Philosophical Transactions of the Royal Society A*, 372(2017).

Khatri, V. and Brown, C.V. (2010) 'Designing data governance', *Communications of the ACM*, 53(1), pp. 148–152.

Leemans, S.J.J., Fahland, D., van der Aalst, W.M.P. (2013), 'Discovering Block-Structured Process Models from Event Logs Containing Infrequent Behaviour'. Business Process Management Workshops. BPM 2013. Lecture Notes in Business Information Processing, vol 171. Springer

Nonaka, I. (1994) 'A dynamic theory of organizational knowledge creation', *Organization Science*, 5(1), pp. 14–37.

Oldenburg, F., Hoberg, K., & Leopold, H. (2025). Process mining in supply chain management: state-of-the-art, use cases and research outlook. International Journal of Production Research, 63(8), 2889-2904.

Ornari, F., Compagnucci, I, De Donato, M.C., Bertrand, Y.,Beyel, H.H., et al. (2025) 'Digital Twins of Business Processes: A research manifesto'. Internet of Things, 30.

Otto, B. (2011) 'Organizing data governance: Findings from the telecommunications industry and consequences for large service providers', *Communications of the Association for Information Systems*, 29(1), pp. 45–66.

Peffers, K., Tuunanen, T., Rothenberger, M.A. and Chatterjee, S. (2007) 'A Design Science Research Methodology for Information Systems Research', *Journal of Management Information Systems*, 24(3), pp. 45–77.

Rojas, E., Munoz-Gama, J., Sepúlveda, M. and Capurro, D. (2016) 'Process mining in healthcare: A literature review', *Journal of Biomedical Informatics*, 61, pp. 224–236.

Ruiz-Garcia, L., Steinberger, G. and Rothmund, M. (2010) 'A model and prototype implementation for tracking and tracing agricultural batch products along the food chain', *Food Control*, 21(2), pp. 112–121.

Striani, F., Colucci, C., Corallo, A., Paiano, R., & Pascarelli, C. (2022). Process Mining in Healthcare: A Systematic Literature Review and A Case Study. Advances in Science, Technology and Engineering Systems Journal , 7(6), 151-160.

van der Aalst, W.M.P. (2016) *Process Mining: Data Science in Action*. Berlin: Springer.

van der Aalst, W.M.P., van Dongen, B.F., Günther, C.W., Rozinat, A., Verbeek, H.M.W. and Weijters, A.J.M.M. (2012) 'Process mining manifesto', in *Business Process Management Workshops*. Berlin: Springer, pp. 169–194.

van Zelst, S.J., Buijs, J.C.A.M., Vázquez-Barreiros, B. et al. (2020), 'Repairing Alignments of Process Models'. Bus Inf Syst Eng 62, pp.289–304

Verdouw, C.N., Wolfert, S., Beulens, A.J.M. and Rialland, A. (2016) 'Virtualization of food supply chains with the Internet of Things', *Journal of Food Engineering*, 176, pp. 128–136.

Weber, K., Otto, B. and Österle, H. (2009) 'One size does not fit all: A contingency approach to data governance', *MIS Quarterly Executive*, 8(4), pp. 207–216.

Yadav, A., Deshpande, P., Kausley, S., Rai, B., Barat, S. (2025) 'Harnessing Digital Twins for Building Resilient Food Supply Chains', Digital Twins for Simulation-Based Decision-Making, Springer, Cham.

From Frugal to Reverse Innovation: A Knowledge-Based Framework for Upstream Diffusion

Faisal Rasool[1], Gustavo Morales-Alonso[2], Vicky O´Rourke[1], Antonio Hidalgo[2]
[1] Atlantic Technological University, Letterkenny, Ireland.
[2] Universidad Politécnica de Madrid, Madrid, Spain.

Abstract

Frugal innovation (FI) has emerged as a strategic response to resource constraints, enabling firms to develop cost-effective solutions that address essential needs in emerging markets. At the same time, reverse innovation (RI) has gained increasing attention as a mechanism through which such solutions diffuse into developed economies. Despite their conceptual proximity, existing literature has largely treated FI and RI as separate domains, leaving underexplored the mechanisms that enable frugal innovations to transition successfully across contexts. This study addresses this gap, drawing on the knowledge-based view (KBV) and dynamic capabilities (DC) perspective. We conceptualise the transition from FI to RI as dependent on two sequential mechanisms: knowledge transformation (transfer readiness) and market validation and translation. A five stages framework is proposed, which allows reframing reverse innovation as a knowledge transformation and validation challenge rather than merely a cost or design issue. A clear processual link between FI capabilities and RI outcomes is established, while introducing the concept of transfer readiness as a key construct explaining why some innovations remain locally embedded while others achieve cross-context diffusion. At the same time, institutional and ecosystem-level factors are integrated into the explanation of adoption outcomes, providing a structured and nuanced understanding of upstream diffusion, offering implications for both theory and practice, and opening avenues for future empirical research.

Keywords: Frugal Innovation, Reverse Innovation, Knowledge-Based View, Dynamic Capabilities, Upstream Diffusion.

Paper type: Academic Research Paper

1 Introduction

Growing economic pressures, sustainability goals, and widening inequalities have forced organizations to deliver more value with fewer resources. While such pressures have historically been associated with emerging economies, they are increasingly evident in advanced markets (Otto, 2025). Within this context, frugal innovation (FI) has emerged as a distinct approach to innovation, characterized by the development of solutions that meet essential performance requirements while significantly reducing cost and resource intensity (Weyrauch & Herstatt, 2017; Zeschky et al., 2011). FI is characterized by its development in emerging markets. For this reason, upstream diffusion requires not only technical adaptation and favorable market conditions but also the restructuring and validation of knowledge.

Additionally, the concept of reverse innovation (RI) is increasingly gaining recognition in explaining how innovations developed in resource-constrained environments diffuse into advanced markets (Govindarajan et al., 2022; Simula et al., 2015). Furthermore, existing research has largely treated FI and RI as related but distinct domains (Rasool et al., 2026). FI literature explains how innovations emerge under constraints (Brem et al., 2020; Hossain, 2021), whereas RI research focuses on diffusion outcomes and strategic implications (Hadengue et al., 2017; Simula et al., 2015). However, the mechanisms that enable frugal innovations to be transferred and adopted in developed markets remain underdeveloped. Recent work suggests that this gap cannot be explained solely by cost or design considerations. Instead, it reflects deeper challenges related to knowledge transformation and legitimacy (Fisac-Garcia et al., 2025).

Although FI and RI are conceptually related (Rasool et al., 2026), the literature still lacks a clear explanation of the mechanisms that enable constraint-born solutions to become transferable and adoptable in developed markets. Many frugal innovations fail to achieve adoption beyond their original context, despite offering clear economic advantages. Previous research has argued that both internal and external factors play a critical role in this transfer and adoption (Hadengue et al., 2017). However, have failed to provide a unified framework for transferring FI developed in developing economies to developed economies.

This study fills this gap by creating a knowledge-based conceptual framework that describes upstream diffusion as a multi-stage process. Using the knowledge-based view (KBV) and dynamic capabilities (DC), we propose that the shift from FI to RI relies on two sequential mechanisms: knowledge transformation (transfer

readiness) and market validation and translation. These mechanisms clarify how locally embedded innovations become transferable, credible, and adoptable in new environments. Accordingly, the study addresses the research question: *How do frugal innovations transition from resource-constrained contexts to adoption in developed markets?* By integrating FI and RI through a process-oriented perspective, this study contributes to the literature by offering a structured explanation of upstream diffusion and identifying the conditions under which it succeeds or fails.

The paper is organized as follows. After this introduction, a review of prior research is summarized in section 2. Section 3 describes the framework development, and section 4 presents an example to illustrate how the proposed framework operates in practice. Finally, section 5 presents the main conclusions and limitations of the study.

2 Theoretical Background

2.1 Frugal Innovation as a Capability

Early work on FI emphasised affordability and "good-enough" performance (Zeschky et al., 2011). Neverthless more recent research, such as by Rossetto et al. (2023) and López-Sánchez & Santos-Vijande (2022) suggests that FI reflects a distinct way of organising innovation activities rather than merely an outcome. Building on this shift, FI can be more accurately conceptualised as a capability, defined by a set of organisational routines that enable firms to systematically innovate under resource constraints (Hossain, 2020; Rosca et al., 2017). At its core, FI capability comprises several interrelated routines. First, constraint-driven problem framing shapes how firms define innovation challenges under scarcity. Rather than starting from technological possibilities, firms begin with limitations such as affordability thresholds and infrastructure gaps, prioritizing essential functionality over feature maximization (Agarwal & Brem, 2017; López-Sánchez & Santos-Vijande, 2022). Second, resource reconfiguration enables firms to construct solutions under constraint by recombining available components, knowledge, and capabilities in novel ways (Cai et al., 2019). This allows firms to translate constraint-defined problems into workable solutions despite limited resources. Third, iterative and feedback-driven development replaces linear innovation processes (Agarwal et al., 2021). FI are often developed through rapid prototyping, experimentation, and continuous user feedback, allowing firms to refine solutions in real-world conditions (Brem et al., 2020; Hossain, 2020). This iterative logic prioritises functionality and usability over technical perfection. Finally, user-centric simplification ensures that solutions focus on essential needs rather than feature complexity (Fisac-Garcia et al., 2025; Rasool et al., 2026). By eliminating non-critical attributes and emphasizing accessibility, firms create offerings that are both affordable and contextually appropriate (Weyrauch & Herstatt, 2017). As a result, these routines offer a capability that enables firms to generate viable solutions under constraints.

2.2 Reverse Innovation and the Challenge of Upstream Diffusion

Upstream diffusion of FIs into developed markets remains limited and inconsistent, despite their effectiveness in resource-constrained contexts (Agarwal & Brem, 2017; Cunha et al., 2014; Fischer et al., 2021). The difficulty lies not FI ability to provide solutions, but in its ability to function and be accepted within fundamentally different context. At the core of this challenge is a contextual mismatch between origin and target markets (Bhatti et al., 2017; Fischer et al., 2021).

FI are typically optimised for affordability, flexibility, and minimal infrastructure dependence. However, innovations envisioned in developed markets are characterised by formal regulatory systems, standardised processes, higher expectations of reliability, and complex procurement structures (Cunha et al., 2014; Von Zedtwitz et al., 2015). As a result, FIs that perform effectively in developing economies may not meet the technical, institutional, or perceptual requirements of higher income markets. Addressing this mismatch requires more than simple diffusion. Several activities and actions need to be taken before the innovation can be implemented in developed economies (Fischer et al., 2021). This may include meeting regulatory standards (Blind, 2016) and customer expectations of quality and functionality (Bhatti et al., 2017). As reported by Wang & Wang (2024) that often fall victim of perception of low quality even when they are technically adequate and perform necessary functions reliably and consistently.

To no surprise, empirical evidence suggests that many FIs fail to achieve upstream diffusion (Fischer et al., 2021). However, this failure is rarely due to a lack of cost advantage or functional performance (Bhatti et al., 2017). Instead, is often a result of firm's inability to bridge the gap between locally embedded innovation practices and globally standardized expectations (Fischer et al., 2021). To address this, firms need to develop

routines and processes that can translate the built-in tacit embedded knowledge in FI into regular innovation targeted towards developed economies. These routines will shift from viewing RI as a diffusion outcome to understanding it as a process shaped by knowledge transformation and contextual validation.

2.3 Knowledge-Based and Dynamic Capabilities Perspectives

The knowledge-based view (KBV) provides a useful starting point by viewing firms in terms of the knowledge they develop and use, and by emphasizing how this knowledge is structured, integrated, and made transferable across contexts . FIs are not merely products or services but are built on knowledge shaped by local constraints, experiential learning, and context-specific problem-solving processes (Fisac-Garcia et al., 2025; Rasool et al., 2026). A key implication of this view is that not all knowledge is equally transferable (Nair et al., 2015). FIs often rely on tacit, informal, and context-dependent knowledge, which is difficult to articulate, codify, and replicate (Felin & Hesterly, 2007; Nonaka & Takeuchi, 1995). While such characteristics enable flexibility and responsiveness, they also limit the extent to which innovations can be understood, evaluated, and replicated beyond their original environment. As a result, FIs that perform effectively in one context may struggle to travel across different institutional and market settings.

This creates a fundamental tension between local effectiveness and cross-context transferability. For RI to occur, the knowledge underlying frugal solutions must be decoupled from its original context and restructured into a form that can be understood, evaluated, and applied in different environments (Felin et al., 2012). However, KBV alone does not fully explain how this transformation takes place. While it highlights how knowledge is structured and embedded, it offers limited insight into how firms actively reconfigure knowledge in response to new demands. To address this limitation, we draw on the Dynamic Capabilities (DC) perspective, which focuses on the processes through which firms sense opportunities, seize them, and reconfigure their resource base in changing environments (Helfat, 2010; Teece, 2007). DCs are particularly relevant, as they explain how firms reconfigure knowledge and capabilities to align with new environments (Helfat, 2010). Specifically, DC enable two critical processes that are central to upstream diffusion. First, firms must develop the ability to transform knowledge by converting tacit, context-specific insights into explicit, structured, and transferable forms (Felin et al., 2012). This involves structuring knowledge so that innovations can be replicated and evaluated beyond their original context. This transformation is not automatic but requires deliberate effort to externalize and organize knowledge in ways that align with formal systems and standards (Nonaka et al., 2000). Second, firms must reconfigure innovations to achieve contextual alignment, ensuring compatibility with regulatory requirements, user expectations, and institutional structures in developed markets (Bhatti et al., 2017). This includes adapting product features, interfaces, and delivery models, as well as generating evidence to support credibility and acceptance.

Building on this, KBV and DC can explain why FIs are difficult to transfer and how firms address this challenge through knowledge transformation and reconfiguration. However, the specific mechanisms linking FI capability to RI outcomes remain under-specified (Rasool et al., 2026). To address this gap, this study develops a mechanism-based framework that conceptualises upstream diffusion as a sequential process involving two stages: knowledge transformation (transfer readiness) and market validation and translation.

3 Framework Development: A Knowledge-Based Mechanism for Upstream Diffusion

Building on the arguments presented in the previous section, this study conceptualizes upstream diffusion as a sequential and mechanism-driven process linking FI capability to RI outcomes. Rather than viewing RI as a simple diffusion outcome, the framework explains it as the result of two critical transformations: knowledge transformation and contextual validation. The framework consists of five stages, connected by two bridging mechanisms, presented in Figure 1. Each stage represents a distinct phase in the evolution of FIs, while the bridging mechanisms explain how innovations move from local embeddedness to global adoption. Importantly, the process is shaped by stage-specific boundary conditions, which influence both the feasibility and effectiveness of these transformations.

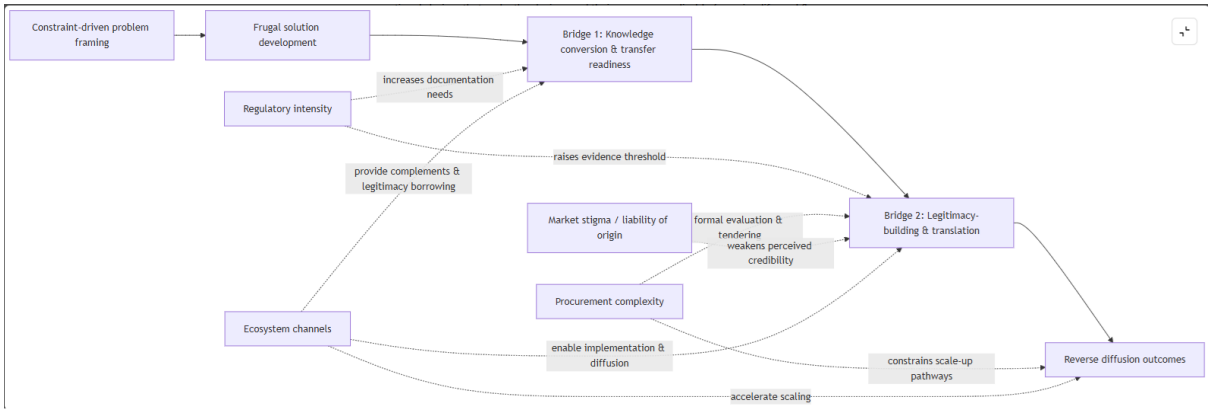


Figure 1: Framework for upstream diffusion

3.1 Stage 1: Constraint-Driven Problem Framing

The process begins with constraint-driven problem framing, where firms define innovation challenges under conditions of scarcity. As discussed earlier, FI capability is rooted in treating constraints such as affordability thresholds, infrastructural limitations, and resource scarcity as design parameters rather than barriers (Agarwal & Brem, 2017; López-Sánchez & Santos-Vijande, 2022). This stage determines the innovation's value logic. Instead of optimizing for performance maximization, firms prioritize essential functionality, accessibility, and cost efficiency. As a result, the innovation is fundamentally shaped by local conditions, embedding assumptions about users, usage environments, and resource availability (Bhatti et al., 2017). However, this localized problem framing introduces the first structural limitation for upstream diffusion. The innovation is not only designed differently but also grounded in context-specific interpretations of value and functionality, which may not align with expectations in developed markets.

3.2 Stage 2: Frugal Solution Development

In the second stage, firms develop solutions by applying FI capability. This involves coordinated routines such as resource reconfiguration, iterative prototyping, and user-centric simplification (Cai et al., 2019; Hossain, 2020). The resulting solution achieves acceptable performance at significantly lower cost, making it highly effective within its original context. However, consistent with the KBV, the knowledge underlying this solution remains largely tacit, experiential, and context dependent (Fisac-Garcia et al., 2025; Rasool et al., 2026). Design decisions are often guided by informal practices and iterative experimentation rather than formalized processes (Felin et al., 2012). At this stage, the innovation is functional but not yet transferable. It lacks the explicit structure required for replication, evaluation, and adoption in different environments. This creates a critical gap between the generation of frugal solutions and their potential for upstream diffusion.

3.3 Bridge 1: Knowledge Transformation (Transfer Readiness)

The first step in moving from locally embedded innovation to a transferable solution is knowledge transformation, referred to here as transfer readiness. Transfer readiness captures how far the knowledge underlying an innovation is structured and made transferable across contexts. It involves turning tacit, context-specific knowledge into explicit and usable forms that can be understood and evaluated beyond the original setting (Felin et al., 2012). From a KBV perspective, this stage represents the externalisation and structuring of knowledge, enabling it to be communicated and evaluated beyond its original context (Felin et al., 2012). From a DCs perspective, it reflects the firm's ability to reconfigure internal routines to align with external requirements (Teece, 2007). This mechanism is a necessary condition for upstream diffusion.

Without transfer readiness, innovations remain tied to their original context, as external actors cannot assess or replicate them. Beyond its core function, knowledge transformation is significantly shaped by regulatory requirements, which determines the level of codification and formalisation required. In highly regulated environments, firms must convert informal and experiential knowledge into externally auditable formats (Blind, 2016), such as technical documentation, compliance reports, and testing protocols. This increases documentation requirements and forces firms to systematically structure their knowledge. As a result, transfer readiness becomes both a technical and institutional construct, shaped by external expectations (Blind, 2016).

3.4 Bridge 2: Legitimacy-Building and Translation

While transfer readiness enables movement across contexts, it does not ensure adoption (Fischer et al., 2021). This stage addresses the challenge of acceptance in developed markets. Legitimacy-building focuses on establishing credibility and reducing uncertainty. Innovations originating from resource-constrained contexts often face scepticism due to perceived quality risks and liability of origin (Bhatti et al., 2017). This creates a credibility gap that must be addressed through activities such as testing, certification, and endorsement (Rao, 1994). Translation, in turn, involves adapting the innovation to align with developed-market expectations. This includes modifying product features, ensuring regulatory compliance, and integrating with existing systems and practices.

The effectiveness of this stage is shaped by market stigma and procurement complexity. Market stigma influences how innovations are perceived, often leading to lower initial credibility (Bhatti et al., 2017). This effect is amplified in formal evaluation settings, where unfamiliar design logics may be interpreted as indicators of lower quality (Fischer et al., 2021). Procurement complexity further constrains adoption by imposing structured evaluation processes that favour established solutions (Uyarra et al., 2014). Firms must therefore ensure that innovations not only perform effectively but also fit within predefined evaluation frameworks. In this context, ecosystem channels play a critical enabling role. Through partnerships with established actors, firms can access complementary capabilities and leverage existing reputational capital, effectively engaging in legitimacy borrowing. This reduces uncertainty and enhances acceptance.

3.5 Stage 5: Reverse Diffusion Outcomes

The final stage focuses on the adoption and scaling of innovations in developed markets, which are shaped by the interaction between ecosystem channels and institutional structures. Ecosystem channels enable implementation and diffusion by providing access to distribution networks, complementary assets, and local market knowledge. They also accelerate scaling by enabling firms to leverage existing infrastructure rather than build new infrastructure (Kapoor, 2018). However, these enabling effects are constrained by procurement complexity, which can limit scale-up pathways (Uyarra et al., 2014). In highly structured systems, even validated innovations may struggle to achieve widespread adoption if they do not align with institutional purchasing mechanisms. As a result, reverse diffusion outcomes depend not only on innovation characteristics but on the alignment between ecosystem support and institutional conditions. Table 1 highlights the stage-wise conditions and failure points in upstream Diffusion.

Table 1**:** Stage-wise conditions and failure points in upstream diffusion

Stage	Key Requirement	If Not Achieved	Outcome
Stage 1–2: Frugal Innovation Development	Functional solution under constraints	Problem–solution misfit	Innovation fails locally
Bridge 1: Knowledge Transformation (Transfer Readiness)	Knowledge structured and transferable	Knowledge remains tacit and embedded	Innovation cannot be transferred
Bridge 2: Market Validation and Translation	Credibility and contextual alignment	Credibility gap or misalignment	Innovation not adopted
Stage 5: Diffusion and Scaling	Alignment with ecosystem and institutions	Structural barriers to scaling	Limited or no diffusion

4 Illustrative Example

To demonstrate how the proposed framework operates in practice, this section presents a stylised illustration based on the development and diffusion of a low-cost portable medical device. Figure 2 summarizes the illustrative example application process. The example is not intended as an empirical case but as a theory-driven application that maps the framework's stages and mechanisms onto a realistic innovation trajectory.

Problem Framing: The innovation originates in a resource-constrained healthcare environment, where access to diagnostic equipment is limited due to high costs, unreliable infrastructure, and a shortage of trained personnel. In this context, the problem is defined in terms of providing accessible and affordable functionality rather than maximizing technical performance. As a result, the innovation team prioritizes portability, low power consumption, and ease of use over advanced features. This reflects a constraint-driven approach, where affordability and usability shape the innovation space.

Solution Development: Building on this framing, the firm develops a portable diagnostic device using simplified components and locally available resources. The development process is iterative, relying on rapid

prototyping and continuous feedback from frontline healthcare workers. The resulting device is cost-effective, robust, and functional within its original context. However, its design logic remains largely implicit, with limited formal documentation or standardization. At this stage, the innovation is effective locally but lacks the structure required for broader diffusion.

Knowledge Transformation: To enter developed markets, the firm must transform the knowledge underlying the device into a form that can be transferred. This involves structuring design and production knowledge so that it can be understood and evaluated beyond the original context. Regulatory requirements play a key role at this stage, as the firm must provide evidence on safety, reliability, and performance. Informal design decisions are translated into formal specifications, and iterative testing is reframed as structured validation. Through these efforts, the device achieves transfer readiness, making it technically transferable across contexts.

Validation and Translation: Despite achieving transfer readiness, the device faces skepticism in developed markets due to its origin and unconventional design. Healthcare providers and regulators question its reliability, creating a credibility gap. To address this, the firm undertakes validation activities such as clinical testing, certification, and pilot implementations (Rao, 1994). At the same time, the device is adapted to align with local expectations, including interface design and system compatibility. Procurement systems pose an additional challenge. Hospitals and public health institutions rely on formal evaluation processes that prioritize standardization and proven track records. These systems may initially weaken perceived credibility, as the device does not fit an established category. To address this, the firm collaborates with established distributors and healthcare allies, leveraging ecosystem channels for access. These partnerships lend credibility, build trust, and ease entry into developed markets.

Diffusion and Scaling: Following successful validation, the device begins to diffuse within cost-sensitive segments of the developed market. Partnerships with established actors provide access to distribution networks and institutional channels, facilitating implementation. These partnerships also support scaling by allowing the firm to leverage existing infrastructure. However, diffusion remains constrained by procurement structures, which limit adoption in more formalized settings. Over time, the device achieves broader uptake where its value proposition aligns with cost and accessibility needs.

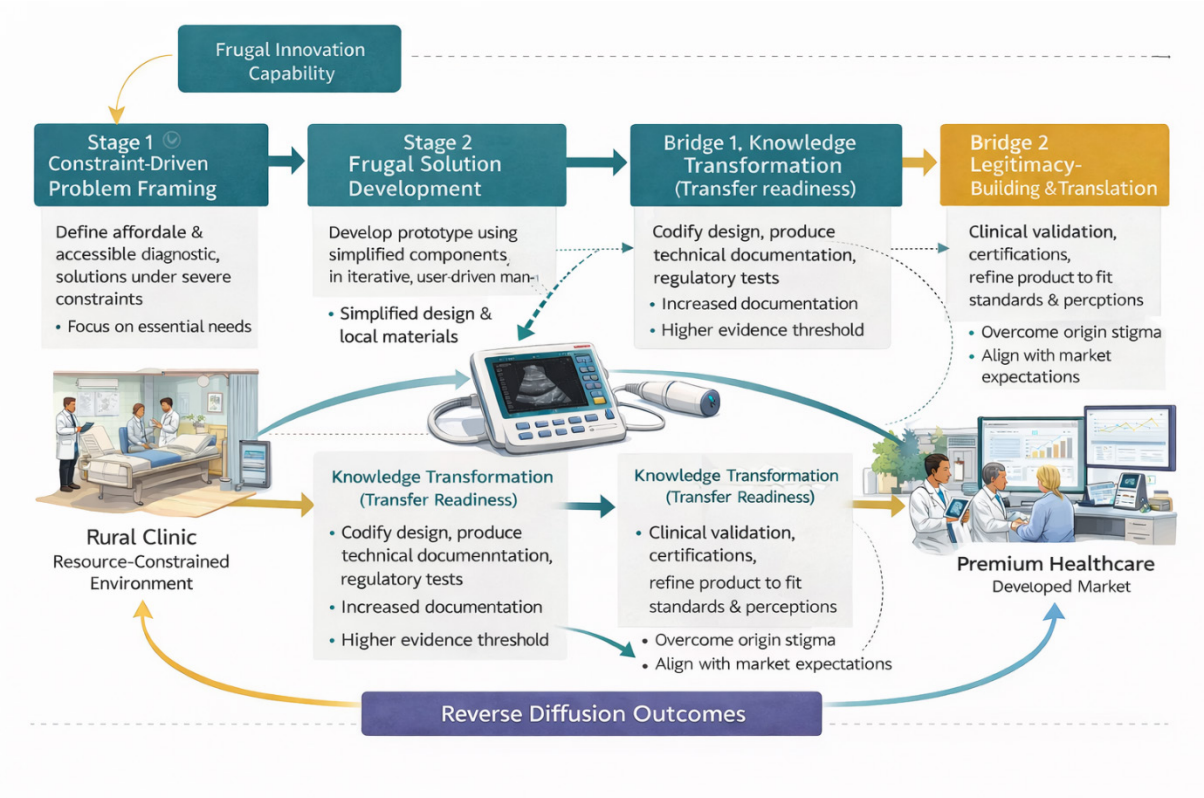


Figure 2: Illustrative example of framework application

5 Conclusions, Implications, and Limitations

This study set out to explain why frugal innovations, despite their demonstrated effectiveness in resource-constrained environments, often fail to achieve adoption in developed markets. By integrating insights from frugal innovation and reverse innovation literatures and grounding the analysis in the knowledge-based view and dynamic capabilities perspective, the study develops a mechanism-based explanation of upstream diffusion. The findings shift the focus from viewing reverse innovation as a diffusion outcome to understanding it as a conditional and staged transformation process, shaped by the structure, transferability, and legitimacy of knowledge.

5.1 Theoretical Contribution

This study makes several contributions to the literature. First, it establishes a clear link between FI capability and RI outcomes. While prior research has examined these domains separately, the connection between them has remained largely implicit. By conceptualising upstream diffusion as a staged process, the study provides a structured explanation of how innovations move from locally embedded solutions to adoption in developed markets. Second, the study reframes RI as a knowledge transformation and validation problem rather than a purely economic or technological one. While cost advantage and market fit remain important, the framework highlights that the primary barrier lies in the nature of the knowledge embedded within frugal innovations. This knowledge is often tacit, context-specific, and embedded in informal practices, which limits its transferability. As a result, successful diffusion depends on the ability to transform this knowledge into structured, evaluable forms and to align it with new contexts. Third, the study introduces transfer readiness as a distinct construct capturing the outcome of knowledge transformation. Unlike related concepts such as absorptive capacity or modularity, transfer readiness focuses on the extent to which internally generated knowledge is structured for external evaluation and replication. Fourth, the study highlights the role of validation and institutional alignment in shaping adoption. Even when innovations are technically transferable, they may fail due to credibility gaps, market stigma, and procurement constraints.

By integrating these factors into a process framework, the study shows that adoption depends not only on technical performance but also on how innovations are evaluated within specific institutional contexts. Finally, the framework explains the heterogeneity observed in reverse innovation outcomes. Rather than attributing success or failure to inherent differences in innovation quality, the study shows that outcomes depend on how effectively firms navigate different stages of the process. Innovations may fail because they are not transferable, not accepted, or not scalable, depending on contextual conditions.

5.2 Implications for Theory and Policy

The findings have important implications for both managers and policymakers. For managers, the framework highlights that success in developed markets requires more than cost-efficient design. Firms must invest in structuring knowledge, generating credible evidence, and aligning innovations with institutional expectations. In particular, developing transfer-ready innovations should be treated as a deliberate capability rather than a by-product of development. Firms must also actively engage in validation and translation activities to ensure that innovations are both credible and contextually aligned.

For policymakers, the findings emphasise the role of regulatory and procurement systems in shaping innovation diffusion. While these systems are necessary for ensuring quality and accountability, excessive complexity can limit the adoption of new solutions. Simplifying certification processes, supporting pilot testing, and fostering collaboration between firms and institutional actors can facilitate the diffusion of frugal innovations in developed markets.

5.3. Limitations and Future Research

While this study offers a structured and mechanism-based explanation of upstream diffusion, several limitations should be acknowledged. First, the framework is conceptual in nature and has not been empirically tested. Although it is grounded in established theoretical perspectives and supported by a theory-driven illustrative example, the proposed relationships and mechanisms require validation through empirical research. Future studies could test the framework using qualitative case studies, longitudinal data, or mixed-method approaches to examine how frugal innovations evolve across different contexts. Second, the framework adopts a firm-level perspective, focusing on how organisations develop, transform, and adapt innovations. While this level of analysis is appropriate for examining capabilities and mechanisms, it may underrepresent broader system-level dynamics. Future research could extend the framework by incorporating ecosystem-level or

institutional perspectives, particularly to explore how policy environments, industry structures, and cross-border collaborations shape reverse innovation outcomes. Third, the concept of transfer readiness, introduced in this study, requires further refinement and operationalisation. While the framework conceptualises it as an outcome of knowledge transformation, additional work is needed to develop measurement approaches and empirical indicators. Future research could examine how transfer readiness varies across industries and how it interacts with other constructs such as absorptive capacity, modularity, or standardisation. Fourth, the framework focuses on a specific pathway of innovation flow from resource-constrained to developed markets. While this captures the essence of reverse innovation, it does not fully account for bidirectional or networked innovation processes, where knowledge flows across multiple contexts simultaneously. Future studies could explore how the proposed mechanisms operate in more complex, multi-directional innovation systems. Finally, while the framework incorporates key boundary conditions such as regulatory intensity, market stigma, procurement complexity, and ecosystem channels, these factors are treated at a conceptual level. Further research is needed to examine how these conditions vary across sectors and regions, and how their interactions influence different stages of the process. Despite these limitations, the framework provides a foundation for future research by offering a mechanism-based and process-oriented perspective on reverse innovation. By identifying key transformation stages and contextual influences, it opens avenues for both theoretical refinement and empirical investigation.

6 Ethics declaration

This research does not need an ethical clearance.

7 AI declaration

In this research, ChatGPT was used to polish the language in the last version of the manuscript.

References

Agarwal, N., & Brem, A. (2017). Frugal innovation-past, present, and future. *IEEE Engineering Management Review*, *45*(3), 37–41. https://doi.org/10.1109/EMR.2017.2734320

Agarwal, N., Oehler, J., & Brem, A. (2021). Constraint-Based Thinking: A Structured Approach for Developing Frugal Innovations. *IEEE Transactions on Engineering Management*, *68*(3), 739–751. https://doi.org/10.1109/TEM.2020.3042929

Bhatti, Y. A., Prime, M., Harris, M., Wadge, H., McQueen, J., Patel, H., Carter, A. W., Parston, G., & Darzi, A. (2017). The search for the holy grail: Frugal innovation in healthcare from low-income or middle-income countries for reverse innovation to developed countries. *BMJ Innovations*, *3*(4), 212–220. https://doi.org/10.1136/bmjinnov-2016-000186

Blind, K. (2016). The impact of regulation on innovation. In *Handbook of Innovation Policy Impact* (pp. 450–482). Edward Elgar Publishing.

Brem, A., Wimschneider, C., De Aguiar Dutra, A. R., Vieira Cubas, A. L., & Ribeiro, R. D. (2020). How to design and construct an innovative frugal product? An empirical examination of a frugal new product development process. *Journal of Cleaner Production*, *275*, 122232. https://doi.org/10.1016/j.jclepro.2020.122232

Cai, Q., Ying, Y., Liu, Y., & Wu, W. (2019). Innovating with Limited Resources: The Antecedents and Consequences of Frugal Innovation. *Sustainability*, *11*(20), 5789. https://doi.org/10.3390/su11205789

Cunha, M. P. E., Rego, A., Oliveira, P., Rosado, P., & Habib, N. (2014). Product Innovation in Resource-Poor Environments: Three Research Streams. *Journal of Product Innovation Management*, *31*(2), 202–210. https://doi.org/10.1111/jpim.12090

Felin, T., Foss, N. J., Heimeriks, K. H., & Madsen, T. L. (2012). Microfoundations of Routines and Capabilities: Individuals, Processes, and Structure. *Journal of Management Studies*, *49*(8), 1351–1374. https://doi.org/10.1111/j.1467-6486.2012.01052.x

Felin, T., & Hesterly, W. S. (2007). The Knowledge-Based View, Nested Heterogeneity, and New Value Creation: Philosophical Considerations on the Locus of Knowledge. *Academy of Management Review*, *32*(1), 195–218. https://doi.org/10.5465/amr.2007.23464020

Fisac-Garcia, R., Rasool, F., Morales-Alonso, G., & Schiuma, G. (2025). Examining the potential of frugal innovation in services: Lessons from an NGO. *European Journal of Innovation Management*, *28*(9), 4806–4826. https://doi.org/10.1108/EJIM-10-2024-1220

Fischer, B., Guerrero, M., Guimón, J., & Schaeffer, P. R. (2021). Knowledge transfer for frugal innovation: Where do entrepreneurial universities stand? *Journal of Knowledge Management*, *25*(2), 360–379. Scopus. https://doi.org/10.1108/JKM-01-2020-0040

Govindarajan, V., Siegel, A., & Trimble, C. (2022). *Reverse innovation: Create far from home, win everywhere* (First edition). Ascent Audio.

Hadengue, M., De Marcellis-Warin, N., & Warin, T. (2017). Reverse innovation: A systematic literature review. *International Journal of Emerging Markets*, *12*(2), 142–182. https://doi.org/10.1108/IJoEM-12-2015-0272

Helfat, C. E. (Ed.). (2010). *Dynamic capabilities: Understanding strategic change in organizations* (Nachdr.). Blackwell Publ.

Hossain, M. (2020). Frugal innovation: Conception, development, diffusion, and outcome. *Journal of Cleaner Production*, *262*, 121456. https://doi.org/10.1016/j.jclepro.2020.121456

Hossain, M. (2021). Frugal innovation: Unveiling the uncomfortable reality. *Technology in Society*, *67*. Scopus. https://doi.org/10.1016/j.techsoc.2021.101759

Kapoor, R. (2018). Ecosystems: Broadening the locus of value creation. *Journal of Organization Design*, *7*(1), 12. https://doi.org/10.1186/s41469-018-0035-4

López-Sánchez, J. Á., & Santos-Vijande, M. L. (2022). Key capabilities for frugal innovation in developed economies: Insights into the current transition towards sustainability. *Sustainability Science*, *17*(1), 191–207. https://doi.org/10.1007/s11625-021-01071-1

Nair, S. R., Demirbag, M., & Mellahi, K. (2015). Reverse Knowledge Transfer from Overseas Acquisitions: A Survey of Indian MNEs. *Management International Review*, *55*(2), 277–301. https://doi.org/10.1007/s11575-015-0242-y

Nonaka, I., & Takeuchi, H. (1995). *The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation* (1st ed). Oxford University Press, Incorporated.

Nonaka, I., Toyama, R., & Konno, N. (2000). SECI, Ba and Leadership: A Unified Model of Dynamic Knowledge Creation. *Long Range Planning*, *33*(1), 5–34. https://doi.org/10.1016/S0024-6301(99)00115-6

Otto, H. E. (2025). From Frugal Solutions to Reverse Innovation: A Knowledge-Based Perspective. In S. Fukushige, T. Nonaka, H. Kobayashi, C. Tokoro, & E. Yamasue (Eds.), *EcoDesign for Circular Value Creation: Volume II* (pp. 187–201). Springer Nature Singapore. https://doi.org/10.1007/978-981-97-9076-0_12

Rao, H. (1994). The Social Construction of Reputation: Certification Contests, Legitimation, and the Survival of Organizations in the American Automobile Industry: 1895–1912. *Strategic Management Journal*, *15*(S1), 29–44. https://doi.org/10.1002/smj.4250150904

Rasool, F., Navarro-Castillo, Y., Morales-Alonso, G., & Hidalgo, A. (2026). Innovating across borders: Insights from frugal and reverse innovation research. *International Journal of Innovation Studies*, S2096248726000093. https://doi.org/10.1016/j.ijis.2026.01.002

Rosca, E., Arnold, M., & Bendul, J. C. (2017). Business models for sustainable innovation – an empirical analysis of frugal products and services. *Journal of Cleaner Production*, *162*, S133–S145. https://doi.org/10.1016/j.jclepro.2016.02.050

Rossetto, D. E., Borini, F. M., Bernardes, R. C., & Frankwick, G. L. (2023). Measuring frugal innovation capabilities: An initial scale proposition. *Technovation*, *121*, 102674. https://doi.org/10.1016/j.technovation.2022.102674

Simula, H., Hossain, M., & Halme, M. (2015). Frugal and reverse innovations – Quo Vadis? *Current Science*, *109*(9), 1567–1572.

Teece, D. J. (2007). Explicating dynamic capabilities: The nature and microfoundations of (sustainable) enterprise performance. *Strategic Management Journal*, *28*(13), 1319–1350. https://doi.org/10.1002/smj.640

Uyarra, E., Edler, J., Garcia-Estevez, J., Georghiou, L., & Yeow, J. (2014). Barriers to innovation through public procurement: A supplier perspective. *Technovation*, *34*(10), 631–645. https://doi.org/10.1016/j.technovation.2014.04.003

Von Zedtwitz, M., Corsi, S., Søberg, P. V., & Frega, R. (2015). A Typology of Reverse Innovation. *Journal of Product Innovation Management*, *32*(1), 12–28. https://doi.org/10.1111/jpim.12181

Weyrauch, T., & Herstatt, C. (2017). What is frugal innovation? Three defining criteria. *Journal of Frugal Innovation*, *2*(1), 1. https://doi.org/10.1186/s40669-016-0005-y

Zeschky, M., Widenmayer, B., & Gassmann, O. (2011). Frugal Innovation in Emerging Markets. *Research-Technology Management*, *54*(4), 38–45. https://doi.org/10.5437/08956308X5404007

Knowledge Transfer and Innovation Spillovers in Aerospace

Antonio Limongiello, Andrea Tomo
University of Naples Federico II, Naples, Italy

Abstract

The Space Economy is one of the most strategic and innovative sectors of the 21st century, characterized by advanced technologies and new market opportunities (Yablonsky, 2018). Its value chain encompasses public and private actors, spanning upstream and downstream activities (Binci et al., 2022). This innovative environment provides a setting for analysing how knowledge generated in aerospace contexts can be transferred to other industries (Smith, 2013). In this sense, Knowledge Management (KM) provides both a process and an organisational perspective through which knowledge integration and transfer can be understood (Cristache et al., 2025). This paper aims to map the KM practices and innovation processes developed in the aerospace industry and to assess their applicability to other industries. Attention is devoted to the mechanisms through which scientific and technical knowledge is translated into innovations capable of generating market impact through new products, services or organisational models. The study adopts a systematic literature review based on the Web of Science database. The search query combined terms related to aerospace, knowledge management and knowledge transfer. The initial search identified 494 papers. After applying the filters "Article & Proceeding Papers", "Management & Business" and "English", the dataset was reduced to 97 publications. The PRISMA protocol ensured transparency and replicability. Papers were excluded when not aligned with KM or when knowledge was treated as a secondary element within engineering-oriented analyses. The final dataset consists of 57 papers. This study develops a framework for qualitative content analysis based on three dimensions (Snyder, 2019): object of transfer (knowledge, technologies, innovation processes), transfer mechanisms (human-relational, organisational, technological) and sectoral proximity (related, distant, unrelated). The objective is to: a) identify points of convergence between aerospace and other industries; b) classify the main strategies of KM (Ardito et al., 2021); c) provide an overview of good practices supporting learning and innovation dynamics across industries.

Keywords: Knowledge Management, Aerospace Industry, Knowledge Transfer, Cross-sector Spillovers, Innovation

Paper type: Academic Research Paper

1 Introduction

The aerospace industry is undergoing a major transformation from mission-oriented public programs toward a New Space Economy (NSE) shaped by commercial, entrepreneurial and multi-actor logics (Mazzucato & Robinson, 2018; Manotti et al., 2025). In this context, space technologies are increasingly viewed not only as strategic assets, but also as enabling platforms for new markets, industrial applications and cross-sector innovation. Such complexity also makes the aerospace industry a relevant setting for studying knowledge processes. In highly interdependent systems, innovation rarely depends on isolated inventions. It often emerges from the integration of distributed expertise, coordinated routines and cumulative learning across multiple organisations and project stages. These conditions make the management and transfer of knowledge particularly critical (Nonaka, 1994). The aerospace sector requires strong system integration capabilities, coordination across suppliers, cooperative R&D and the management of complex product architectures with many interdependent components (Kim et al., 2022). Collaboration networks are also hierarchical and asymmetric. This affects the distribution of innovation opportunities and access to strategic capabilities (Biggiero & Angelini, 2015). In addition, the innovative performance of space firms is influenced by cluster configuration, technological specialization and interactions with public and private actors in the ecosystem (Teirlinck & Khoshnevis, 2020). Existing studies have already examined important parts of this phenomenon. Some focus on innovation effects generated by public procurement and public-private collaboration. Others analyse inter-organisational transfer frameworks and mechanisms of knowledge movement within and across organisations (Battistella et al., 2016; Castelnovo et al., 2023; Argote et al., 2022). This study addresses the following research questions:

- ***RQ1.*** *What types of knowledge, technologies and innovation processes are transferred or implemented across aerospace and non-aerospace contexts?*
- ***RQ2.*** *Through which human, organisational, relational and technological mechanisms do transfer and implementation processes occur?*
- ***RQ3.*** *How does sectoral proximity shape the transferability and recombination of aerospace knowledge?*

- ***RQ4.*** *How do opportunities and threats associated with transfer and implementation processes affect innovation outcomes and new market applications?*

This study contributes to the knowledge transfer literature by integrating three dimensions (object, mechanism and sectoral proximity) into a unified analytical framework. While prior studies have typically examined these dimensions in isolation (Argote et al., 2022; Battistella et al., 2016; Alberti & Pizzurno, 2015), this research demonstrates that knowledge transfer processes are better understood as multidimensional configurations emerging from their interaction.

2 Theoretical Background

From a KM perspective, competitive advantage does not depend solely on the possession of technological assets but also on the organisational capability to create, integrate, share and apply knowledge effectively across different contexts and levels of analysis (Nonaka, 1994; Argote et al., 2022). The aerospace sector represents a particularly relevant empirical setting because it combines technological complexity, high reliability requirements, cumulative learning processes and advanced problem-solving routines, making it a fertile context for observing how valuable knowledge is generated and redeployed beyond its original boundaries (Teissandier et al., 2026; Wittenborg et al., 2025). Two core concepts underpin this study: knowledge transfer and knowledge implementation (Nonaka, 1994; Argote & Ingram, 2000; Szulanski, 1996). These concepts are analytically distinct yet strongly interconnected within organisational learning and innovation processes. Existing knowledge assets become strategically relevant when organisations implement, transfer and redeploy them across organisational and sectoral boundaries (Okwechime, 2025; Argote & Ingram, 2000). Implementation may involve the replication, adaptation and enactment of routines across organisational contexts (D'Adderio, 2014). The effectiveness of implementation depends on the degree of alignment between new routines and the recipient organisational context (Bertels et al., 2016). Consistent with broader knowledge-to-action perspectives, implementation is strongly influenced by prior problem framing, contextual analysis and the selection of appropriate transfer interventions rather than by adoption alone (Ward et al., 2009). Knowledge transfer is commonly defined as the process through which one organisational unit is influenced by the experience of another (Argote & Ingram, 2000). From an inter-organisational perspective, transfer can be understood as an intentional and goal-oriented interaction between distinct entities through which technological and organisational knowledge is exchanged, recombined or expanded (Battistella et al., 2016).

3 Research Method

This article proposes a systematic literature review using Web of Science. The initial search yielded 494 papers, of which 397 articles were excluded after the application of the following filters: ("Articles & Proceeding Papers", "Management & Business", "English").

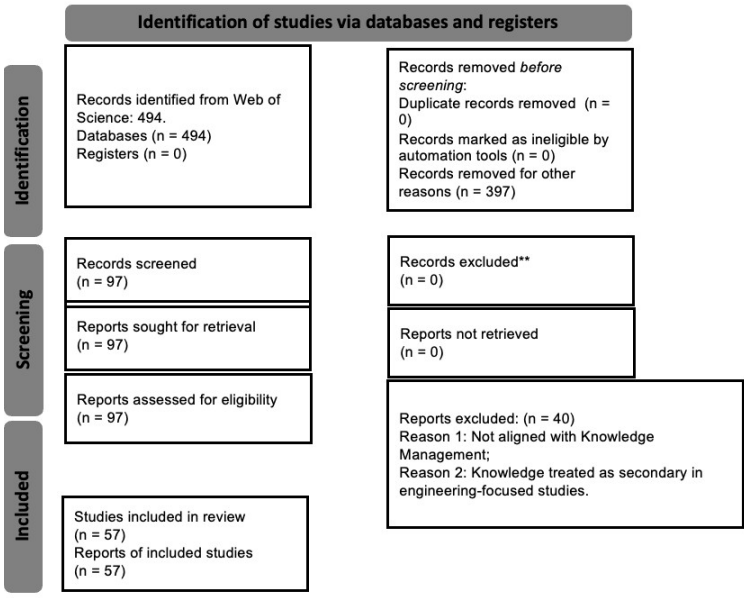


Figure 1. Identification of studies
Source: Page et al., 2021.

In methodological terms, the study follows the logic proposed by Snyder (2019), according to which literature reviews should not only collect prior evidence but also synthesize fragmented knowledge into clearer conceptual structures. In substantive terms, the coding framework draws on key contributions in the knowledge transfer literature. Argote et al. (2022) identify mechanisms and components that shape transfer processes, while Argote (2024) further highlights the importance of knowledge characteristics, motivation and recipients' consideration in explaining transfer outcomes. Battistella et al. (2016) provide an inter-organisational perspective centred on actors, object, channels, relationships and contextual conditions. In this study, "transfer" refers to knowledge flows from aerospace to non-aerospace sectors, while "implementation" captures the deployment or adaptation of knowledge within aerospace contexts or in closely related sectors whose practices may inform aerospace applications. The inclusion of implementation cases beyond direct aerospace transfer reflects an analytical choice aimed at capturing transferable knowledge practices from cognitively or industrially proximate sectors. These cases are not treated as direct spillovers, but as informative analogues that may support knowledge recombination within aerospace contexts.

4 Findings

The analytical framework was derived from the literature and further refined through the authors' interpretative synthesis. It was then used to systematically organise the findings emerging from the 57 reviewed papers across six analytical dimensions, including the three core dimensions of object, mechanism and sectoral proximity. These dimensions were defined in accordance with the research objectives and KM perspective. Specifically, each paper was examined through this pre-defined framework (Figure 2). Due to conference length constraints, the full article-by-article coding table derived from the analysis of the 57 reviewed studies is not reproduced in this version of the paper. In the extended manuscript, this material is reported through detailed appendix tables that systematically classify each contribution across the analytical dimensions adopted in this study. The findings discussed below therefore represent a synthetic cross-case interpretation grounded in that complete coding structure, rather than standalone descriptive observations. This choice was made to improve readability and focus the conference paper on the main patterns, theoretical implications and managerial relevance emerging from the review.

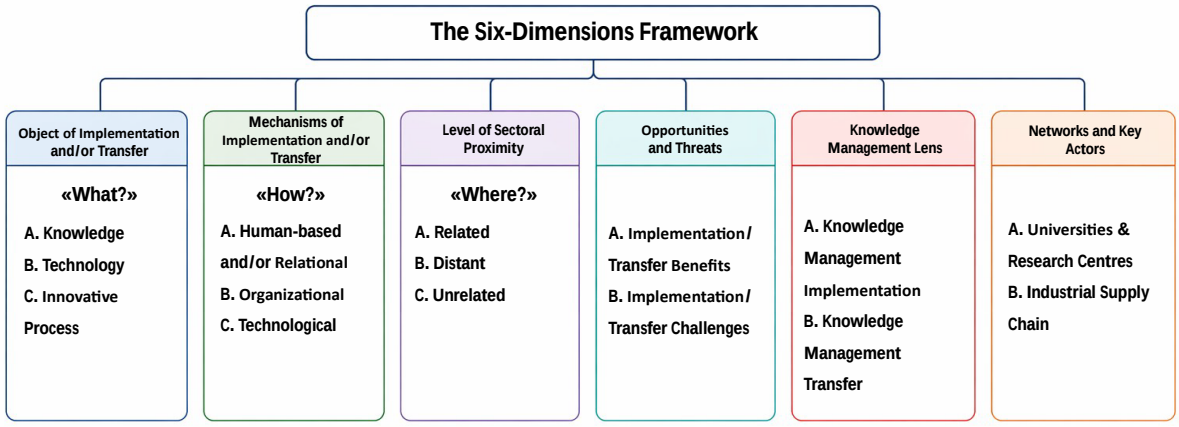


Figure 2. The six-dimensions framework
Source: Authors' own work.

4.1 Object of Implementation and/or Transfer

Across the analysed studies, three recurring categories emerge: knowledge, technology and innovative processes. Knowledge-related subjects include engineering expertise, managerial capabilities, market knowledge and organisational routines (e.g., Smith, 2013; Chirumalla, 2013; Philbin, 2010). Technological subjects involve, for instance, digital systems, hardware and software modules, monitoring infrastructures or Industry 4.0 technologies (e.g., Grubic & Peppard, 2016; Oshri & Newell, 2005; Dyba & De Marchi, 2022). Process-oriented subjects concern servitization models, collaborative innovation routines, new governance approaches and capability-development practices (e.g., Leoni, 2015; Buganza & Verganti, 2009; Johansson, 2014; Corallo et al., 2010). This suggests that cross-sector innovation is not limited to the movement of artefacts but also includes intangible and organisational assets.

Finding 1. The evidence suggests that what is transferred or implemented across sectors cannot be reduced to a single category resource. Rather, the analytical framework developed in this study, drawing on prior

literature and refined through the authors' interpretative synthesis, identifies three main objects: knowledge, technology and innovation processes. Knowledge includes: a) engineering knowledge (n = 15); b) organisational capabilities (n = 14); c) market-oriented understanding (n = 5). Technology refers to: a) digital systems, hardware, software modules, infrastructures and advanced technical solutions (n = 8). Innovation processes concern new ways of: a) organizing activities, coordinating actors, developing services (n = 32); b) applying technologies in novel settings (n = 4). The findings show that cross-sector transfer mainly concerns organizing, coordinating and service-development practices (n = 32), suggesting that innovation often derives from managerial and process reconfiguration rather than from isolated technologies. Engineering knowledge (n = 15) and organisational capabilities (n = 14) also emerge as complementary assets, indicating that technical expertise requires coordination and learning routines to create value. By contrast, standalone technologies (n = 8), market-oriented understanding (n = 5) and novel applications (n = 4) appear less frequent.

4.2 Mechanisms of Implementation and/or Transfer

The evidence suggests that implementation and transfer are enabled through multiple and often complementary mechanisms. Several papers emphasise human and relational mechanisms, such as specialised technicians, joint researchers, collaborative teams, and brokerage roles connecting otherwise separated actors (e.g., Israilidis et al., 2015; Holford, 2010). Others highlight organisational mechanisms, including open innovation agreements, R&D partnerships, university-industry collaboration and supply-chain coordination (e.g., Braziotis & Tannock, 2011; Ehret & Cooke, 2010; Esangbedo et al., 2024). A third group points to technological mechanisms, such as component sharing, digital platforms, Web-based systems, and advanced communication infrastructures (e.g., Oshri & Newell, 2005; Corallo et al., 2012; Grubic & Peppard, 2016).

Finding 2. Transfer and implementation processes do not occur through a single channel. Instead, the application of the framework to the reviewed studies highlights three recurrent mechanisms: human-based & relational, organisational and technological. These mechanisms explain how knowledge, technologies and innovation processes are mobilised, adapted and made operational across contexts. The numerical identifiers reported below indicate the studies associated with each mechanism: a) human-based and relational (n = 46); b) organisational (n = 37); c) technological (n = 29). Human-based and relational mechanisms are the most recurrent channel of transfer and implementation (n = 46), highlighting the centrality of collaboration, trust, mobility and network ties in knowledge circulation. Organisational mechanisms (n = 37) also play a key role through routines, governance and coordination practices that sustain transfer over time. Technological mechanisms are less frequent (n = 29), suggesting that technology alone is insufficient.

4.3 Level of Sectoral Proximity

The studies reveal heterogeneous patterns of sectoral proximity. Some transfers occur toward related sectors where technological and cognitive similarities reduce adaptation costs (e.g., Smith, 2013; Grubic & Peppard, 2016). Other contributions involve more distant or unrelated sectors, where aerospace-derived capabilities are recombined to generate novel applications (e.g., Su et al., 2007). In further cases, the papers do not describe an explicit intersectoral transfer but analyse implementation processes within other industries that may offer lessons transferable to aerospace contexts (e.g., Alberti & Pizzurno, 2015; Chang et al., 2012).

Finding 3. The findings show that sectoral proximity is not a binary condition but a continuum that depends on the nature of the knowledge involved, the recipient context and the KM lens adopted in the analysis. In some studies, the phenomenon concerns an effective transfer from the aerospace sector to non-aerospace domains. In other cases, the evidence refers to implementation processes rather than direct intersectoral transfer. Accordingly, sectoral proximity was used to distinguish between related, distant or unrelated sectors, while also capturing whether the observed contribution concerns transfer or implementation: a) related (n = 41); b) distant (n = 12), c) unrelated (n = 5). The literature is mainly concentrated in related sectors (n = 41), with fewer cases in distant (n = 12) and unrelated sectors (n = 5). This suggests that sectoral proximity facilitates transfer by reducing adaptation costs and easing collaboration through shared knowledge bases and similar requirements.

4.4 Opportunities and Threats

The literature highlights a dual dynamic in which opportunities and threats coexist. On the opportunity side, implementation and transfer processes are associated with stronger innovation capacity, access to new markets, efficiency gains, enhanced competitiveness and the creation of new collaborative ecosystems (e.g., Philbin, 2010; Smith, 2013; Leoni, 2015). For example, cluster strength, network embeddedness and cooperation with educational institutions are presented as important levers for value creation (e.g., Alberti & Pizzurno, 2015;

Esangbedo et al., 2024). On the threat side, authors report risks linked to imitation by competitors, weak coordination, communication barriers, organisational resistance, capability gaps and difficulties in integrating heterogeneous knowledge bases (e.g., Mey & McMahon, 2009; Jafari et al., 2011; Esangbedo et al., 2024).

Finding 4.

Macro-theme	Count	% on filled cells
Knowledge Management & Knowledge Flows	11	33.3
Technology Enablement	6	18.2
Networks, Clusters & Collaboration	5	15.2
Performance, Efficiency & Competitiveness	4	12.1
Innovation & New Development	2	6.1
Skills & Human Capital	2	6.1
Security & Defence Improvement	2	6.1
Risk, Learning & Resilience	1	3.0

Figure 3: Opportunity Macro-themes; n = 33 filled cells
Source: Authors' own work.

The main opportunities identified concern knowledge management and knowledge flows (33,3%), showing that transfer creates value by improving how knowledge is organised, shared and mobilised across boundaries. Other relevant opportunities involve technology enablement (18,2%) and networks, clusters and collaboration (15,2%), which support new applications and inter-organisational exchange. Additional benefits include higher performance, efficiency and competitiveness (12,1%), as well as innovation, skills development, resilience and long-term adaptive capacity.

Macro-theme	Count	% on filled cells
Cultural, Organizational & Communication Barriers	8	24.2
Monitoring, Cost & Managerial Burden	8	24.2
Security, Leakage & Information Protection	6	18.2
Adaptation & Compatibility Challenges	5	15.2
Knowledge Loss & Tacit Dependency	2	6.1
Risk, Safety & Operational Complexity	2	6.1
Training & Capability Gaps	1	3.0
Competitive Pressure	1	3.0

Figure 4: Threat Macro-themes; n = 33 filled cells
Source: Authors' own work.

The main threats concern cultural, organisational and communication barriers, as well as monitoring, cost and managerial burden (24,2% each), showing that transfer is often hindered by coordination problems and implementation complexity. Additional risks include security and knowledge leakage (18,2%), adaptation challenges (15,2%), tacit knowledge loss, operational risks, capability gaps and competitive pressure, all of which can weaken successful transfer outcomes.

4.5 Knowledge Management Lens

Most studies can be interpreted through two dominant KM perspectives: knowledge transfer and knowledge implementation. The first perspective is visible when the emphasis is placed on moving expertise, technologies or routines across organisational and sectoral boundaries (e.g., Smith, 2013; Corallo et al., 2016; Greco et al., 2025). The second emerges when the focus concerns internal dissemination, collaborative learning and the collective use of distributed knowledge (e.g., Davenport & Prusak, 1998; Grimaldi et al., 2015; Duffield & Whitty, 2015).

Finding 5. The findings show a prevalence of an internal KM perspective (n = 39) over an external one (n = 18). Here, the external lens refers to transfer processes, while the internal lens captures implementation

dynamics, including aerospace-to-aerospace cases or practices developed in other sector but potentially applicable to aerospace.

4.6 Networks and Key Actors

Finally, the synthesis confirms the centrality of networks and multi-actor configurations. Universities, research centres, industrial supply chains, startups, public institutions and intermediary organisations recurrently appear as pivotal actors in enabling implementation and transfer (e.g., Buganza & Verganti, 2009; Latino et al., 2016; Greco et al., 2025). Some studies explicitly stress the role of clusters as coordination platforms (e.g., Wu et al., 2015; Grimaldi et al., 2012), while others underline the importance of third-party brokers, strategic alliances or collaborative R&D programmes (e.g., Philbin, 2010; Alberti & Pizzurno, 2017).

Finding 6. The findings show that high relevance is mainly associated with the industrial supply chain (n = 17), followed by universities (n = 12) and research centres (n = 6). This suggests that knowledge transfer and implementation processes are primarily embedded in industrial networks, with academic and research actors playing a supportive but relevant role.

5 Discussion

The findings suggest that aerospace spillovers should not be interpreted as simple one-directional transfers of isolated technologies from one sector to another. Rather, they emerge as multidimensional and non-linear phenomena shaped by the interaction between different types of resources, transfer mechanisms, contextual conditions and actor configurations (Corallo et al., 2012; Smith, 2013; Parihar & Sidharth, 2025). The diffusion of knowledge, technologies and innovation processes depends on whether they can be translated, adapted, recombined and embedded in new organisational settings. This perspective moves beyond deterministic views of spillovers and supports a more systemic understanding of cross-sector innovation dynamics (Argote, 2024). A second theoretical implication concerns the relative importance of mechanisms over artefacts. While prior discussions often focus on what is transferred, the review indicates that successful outcomes depend more strongly on how transfer occurs. The same technological asset may generate very different results depending on the presence of collaborative routines, absorptive capabilities, inter-organisational trust, mobility of skilled personnel or structured interfaces between actors (Grimaldi et al., 2015; Pérez-Valls et al., 2019; Ferreira et al., 2022). The recurrent role of universities, research centres, industrial partners, startups, clusters and intermediaries highlights that aerospace spillovers are deeply embedded in networks rather than produced by isolated organisations. Startups often act as agile recombiners of existing technologies into new applications (Philbin, 2010; Alberti & Pizzurno, 2017; Dalmarco et al., 2019; Dyba & De Marchi, 2022). However, proximity should not be considered a sufficient or deterministic condition. Several studies indicate that distant sectors may generate significant innovation precisely because they combine aerospace knowledge with unrelated competences, needs and market logics (Grubic & Peppard, 2016).

6 Conclusion

This study contributes to the knowledge transfer and KM literature by developing an integrated framework based on object, mechanisms and sectoral proximity. The framework extends prior fragmented approaches by showing that aerospace spillovers are configurational phenomena shaped by the interaction among multiple dimensions rather than linear transfers of isolated technologies. In addition, the aerospace industry emerges as a valuable empirical setting for understanding how complex knowledge assets can be redeployed across sectoral boundaries. The framework, developed through a systematic literature review approach (Wittenborg et al., 2025; Snyder, 2019), will be validated through empirical cases designed to demonstrate both the robustness of the reported relationships and the replicability/generalizability of the proposed model (Fauri, 2025). In this sense, the framework is intended as an analytical device able to specify what has been transferred, how the transfer has occurred, between which actors, which empirical cases substantiate the process, and which forms of innovation and/or sectoral convergence have emerged because of transfer dynamics. These considerations also generate several managerial and policy implications. More broadly, the framework may be conceived as a set of guiding principles for good practices to be adopted by knowledge workers (Ferreira et al., 2022) and, more generally, by specialized intermediaries operating in knowledge and technology transfer processes, such as universities, research centres and consulting organisations (Greco et al., 2025; Esangbedo et al., 2024; Pereira et al., 2021). From a managerial and policy perspective, investments should not focus only on technological assets, but also on relational mechanisms, collaborative routines, intermediary actors and ecosystem architectures that enable knowledge diffusion. Firms may benefit from exploring both related and cognitively

distant sectors, while public institutions may foster cross-sector partnerships and broader socio-economic returns from aerospace capabilities. This study is limited by its reliance on secondary data and the exclusion of engineering and geopolitical contributions. Future research should extend the review and develop case studies to empirically test the proposed framework.

7 Ethics Declaration

This study is based exclusively on a systematic literature review of publicly available academic sources indexed in Web of Science. No human participants, personal data, interviews, surveys, experiments, or confidential organisational data were involved. Therefore, formal ethical clearance was not required for this research according to standard institutional procedures for secondary-source review studies.

8 AI Declaration

No AI tools were used to generate the submitted research content. The manuscript, including its conceptual development, literature analysis, interpretation of findings, and final written version, was independently developed and finalized by the authors.

References

Alberti, F. G., & Pizzurno, E. (2015). Knowledge exchanges in innovation networks: evidences from an Italian aerospace cluster. Competitiveness review, 25(3), 258-287.

Alberti, F. G., & Pizzurno, E. (2017). Oops, I did it again! Knowledge leaks in open innovation networks with start-ups. European journal of innovation management, 20(1), 50-79.

Ardito, L., Natalicchio, A., Appio, F. P., & Petruzzelli, A. M. (2021). The role of scientific knowledge within inventing teams and the moderating effects of team internationalization and team experience: Empirical tests into the aerospace sector. Journal of Business Research, 128, 701-710.

Argote, L. (2024). Knowledge transfer within organisations: Mechanisms, motivation, and consideration. *Annual review of psychology*, *75*(1), 405-431.

Argote, L., & Ingram, P. (2000). Knowledge transfer: A basis for competitive advantage in firms. *Organizational behavior and human decision processes*, *82*(1), 150-169.

Argote, L., Guo, J., Park, S. S., & Hahl, O. (2022). The mechanisms and components of knowledge transfer: The virtual special issue on knowledge transfer within organisations. Organization Science, 33(3), 1232-1249.

Battistella, C., De Toni, A. F., & Pillon, R. (2016). Inter-organisational technology/knowledge transfer: a framework from critical literature review. *The Journal of Technology Transfer*, *41*(5), 1195-1234.

Bertels, S., Howard-Grenville, J., & Pek, S. (2016). Cultural molding, shielding, and shoring at Oilco: The role of culture in the integration of routines. *Organization Science*, *27*(3), 573-593.

Biggiero, L., & Angelini, P. P. (2015). Hunting scale-free properties in R&D collaboration networks: Self-organisation, power-law and policy issues in the European aerospace research area. Technological Forecasting and Social Change, 94, 21-43.

Binci, D., Appolloni, A., & Cheng, W. (2022). Space Economy, Open Innovation and Sustainability. *Symphonya*, (2), 110-118.

Braziotis, C., & Tannock, J. (2011). Building the extended enterprise: key collaboration factors. The International Journal of Logistics Management, 22(3), 349-372.

Buganza, T., & Verganti, R. (2009). Open innovation process to inbound knowledge: Collaboration with universities in four leading firms. European Journal of Innovation Management, 12(3), 306-325.

Castelnovo, P., Clò, S., & Florio, M. (2023). A quasi-experimental design to assess the innovative impact of public procurement: An application to the Italian space industry. Technovation, 121, 102683.

Chang, C. M., Hsu, M. H., & Yen, C. H. (2012). Factors affecting knowledge management success: the fit perspective. Journal of Knowledge Management, 16(6), 847-861.

Chirumalla, K. (2013). Managing knowledge for product-service system innovation: the role of Web 2.0 technologies. Research-Technology Management, 56(2), 45-53.

Corallo, A., Lazoi, M., & Secundo, G. (2012). Inter-organisational knowledge integration in collaborative NPD projects: Evidence from the aerospace industry. Knowledge Management Research & Practice, 10(4), 354-367.

Corallo, A., Lazoi, M., Margherita, A., & Scalvenzi, M. (2010). Optimizing competence management processes: a case study in the aerospace industry. Business Process Management Journal, 16(2), 297-314.

Corallo, A., Lazoi, M., Secundo, G., & DePaolis, P. (2016). An interpretative model from the elasticity theory to explore knowledge integration in new product development. *Knowledge Management Research & Practice*, *14*(4), 478-488.

Cristache, N., Croitoru, G., & Florea, N. V. (2025). The influence of knowledge management on innovation and organisational performance. *Journal of Innovation & Knowledge*, *10*(5), 100793.

D'Adderio, L. (2014). The replication dilemma unravelled: How organisations enact multiple goals in routine transfer. *Organization Science*, *25*(5), 1325-1350.

Davenport, T. H., & Prusak, L. (1998). Working knowledge: How organisations manage what they know. Harvard Business Press.

Duffield, S., & Whitty, S. J. (2015). Developing a systemic lessons learned knowledge model for organisational learning through projects. International journal of project management, 33(2), 311-324.

Dyba, W., & De Marchi, V. (2022). On the road to Industry 4.0 in manufacturing clusters: the role of business support organisations. Competitiveness Review: An International Business Journal, 32(5), 760-776.

Ehret, O., & Cooke, P. (2010). Conceptualising aerospace outsourcing: Airbus UK and the lean supply approach. International Journal of Technology Management, 50(3/4), 300-317.

Esangbedo, C. O., Zhang, J., Esangbedo, M. O., Kone, S. D., & Xu, L. (2024). The role of industry-academia collaboration in enhancing educational opportunities and outcomes under the digital driven Industry 4.0. Journal of Infrastructure, Policy and Development, 8(1), 2569.

Fauri, F. (2025). From Railways to Aircraft: Officine Meccaniche Reggiane's Successful Product Transition in the 1930s. Enterprise & Society, 26(1), 170-196.

Ferreira, J. J., Fernandes, C. I., Guo, Y., & Rammal, H. G. (2022). Knowledge worker mobility and knowledge management in MNEs: A bibliometric analysis and research agenda. Journal of Business Research, 142, 464-475.

Greco, M., Mignacca, B., Locatelli, G., Grimaldi, M., & Rogo, F. (2025). Can you tell a good open innovation failure from a bad one?. *Journal of Knowledge Management*, *29*(10), 3360-3378.

Grimaldi, M., Cricelli, L., & Rogo, F. (2012). A methodology to assess value creation in communities of innovation. Journal of Intellectual Capital, 13(3), 305-330.

Grimaldi, M., Cricelli, L., & Rogo, F. (2015). Assessment and management of intellectual capital: a single case study on the construction and implementation of an IC index. *Knowledge and Process Management*, *22*(4), 235-249.

Grubic, T., & Peppard, J. (2016). Servitized manufacturing firms competing through remote monitoring technology: An exploratory study. Journal of Manufacturing Technology Management, 27(2), 154-184.

Holford, W. D. (2010). Knowledge construction and risk induction/mitigation in dialogical workgroup processes. Qualitative Research in Organisations and Management: An International Journal, 5(2), 127-161.

Israilidis, J., Siachou, E., Cooke, L., & Lock, R. (2015). Individual variables with an impact on Knowledge Transfer: the critical role of employees' ignorance. Journal of knowledge management, 19(6), 1109-1123.

Jafari, M., Akhavan, P., & Nourizadeh, M. (2013). Classification of human resources based on measurement of tacit knowledge: An empirical study in Iran. Journal of Management Development, 32(4), 376-403.

Jafari, M., Rezaeenour, J., Mahdavi Mazdeh, M., & Hooshmandi, A. (2011). Development and evaluation of a knowledge risk management model for project-based organisations: A multi-stage study. Management Decision, 49(3), 309-329.

Johansson, C. (2014). Managing uncertainty and ambiguity in gates: Decision making in aerospace product development. International Journal of Innovation and Technology Management, 11(02), 1450012.

Kim, B., Cho, Y., & Yang, M. G. (2019). Technology acquisition strategy: A latecomer's perspective on integrating component suppliers with system integrators. IEEE Transactions on Engineering Management, 69(6), 2572-2587.

Latino, M. E., Corallo, A., Capone, I., Martino, D., & Trifoglio, A. (2016). Lesson learned and best practice management: A tool to support the enterprise. Knowledge and Process Management, 23(3), 230-244.

Leoni, L. (2015). Adding service means adding knowledge: an inductive single-case study. Business Process Management Journal, 21(3), 610-627.

Manotti, J., Sanasi, S., & Ghezzi, A. (2025). Sustainable business model innovation: A technology affordance perspective in the New Space Economy. Technovation, 143, 103226.

Mazzucato, M., & Robinson, D. K. (2018). Co-creating and directing Innovation Ecosystems? NASA's changing approach to public-private partnerships in low-earth orbit. Technological Forecasting and Social Change, 136, 166-177.

Nonaka, I. (1994). A dynamic theory of organisational knowledge creation. *Organization science*, *5*(1), 14-37.

Okwechime, E. (2025). Business in the Space Economy: Ecosystems, Stakeholders, and the Adoption of the New Space Paradigm in the United Kingdom. *New Space*, *13*(3), 74-76.

Oshri, I., & Newell, S. (2005). Component sharing in complex products and systems: challenges, solutions, and practical implications. IEEE transactions on engineering management, 52(4), 509-521.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... & Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *bmj*, *372*.

Parihar, Y., & Sidharth, S. (2025). Strategic knowledge management framework for helicopter operations (SKMF-HO) to enhance safety and performance. Journal of Advances in Management Research, 1-37.

Pereira, B., Lohmann, G., & Houghton, L. (2021). The role of collaboration in innovation and value creation in the aviation industry. Journal of Creating Value, 7(1), 44-59.

Pérez-Valls, M., Céspedes-Lorente, J., Martínez-del-Río, J., & Antolín-López, R. (2019). How organisational structure affects ecological responsiveness. Business & Society, 58(8), 1634-1670.

Philbin, S. P. (2010). Developing and Managing University-Industry Research Collaborations through a Process Methodology/Industrial Sector Approach. Journal of Research Administration, 41(3), 51-68.

Smith, D. J. (2013). Power-by-the-hour: The role of technology in reshaping business strategy at Rolls-Royce. *Technology analysis & strategic management*, *25*(8), 987-1007.

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. Journal of business research, 104, 333-339.

Su, N. M., Wilensky, H., Redmiles, D., & Mark, G. (2007, November). The gospel of knowledge management in and out of a professional community. In Proceedings of the 2007 ACM International Conference on Supporting Group Work (pp. 197-206).

Szulanski, G. (1996). Exploring internal stickiness: Impediments to the transfer of best practice within the firm. *Strategic management journal*, *17*(S2), 27-43.

Teirlinck, P., & Khoshnevis, P. (2020). Within-cluster determinants of output efficiency of R&D in the space industry. Omega, 94, 102039.

Teissandier, M., Chebbi, H., & Ben Selma, M. (2026). Knowledge management in constraining contexts: applying microfoundations to overcome knowledge stickiness and stimulate innovation. *Journal of Knowledge Management*, 1-23.

Ward, V., House, A., & Hamer, S. (2009). Developing a framework for transferring knowledge into action: a thematic analysis of the literature. *Journal of health services research & policy*, *14*(3), 156-164.

Wittenborg, T., Baimuratov, I., Franzén, L. K., Staack, I., Römer, U., & Auer, S. (2025). Knowledge-Based Aerospace Engineering--A Systematic Literature Review. *arXiv preprint arXiv:2505.10142*.

Wu, W., Yu, B., & Spender, J. C. (2015). Domains and opportunities in knowledge and aerospace management in China: an integrative perspective. Chinese Management Studies, 9(4), 473-481.

Yablonsky, S. (2018). A multidimensional framework for digital platform innovation and management: from business to technological platforms. *Systems Research and Behavioral Science*, *35*(4), 485-501.

Climate Change Adaptation and Mitigation through Digital Technologies in the Extra Virgin Olive Oil Supply Chain

Giulia Padovani, Dario Barberini, Maria Giovina Pasca, Gabriella Arcese
Niccolò Cusano University, Rome, Italy

Abstract

The study aims to analyse the role of digital technologies across the extra virgin olive oil (EVOO) supply chain, with particular reference to their contributions to climate change adaptation and mitigation. The research is developed through a systematic literature review, following the approach proposed by Tranfield et al. (2003), to examine the existing scientific literature on the topic in a structured manner. The analysis was conducted in January 2026 using the Scopus database, applying specific inclusion criteria for document type, language, and full-text availability. The final corpus consists of 20 articles, selected through a full-text screening and evaluation process conducted by two researchers.

The results focus on three main themes: I) digitalisation and operational optimisation in olive cultivation; II) predictive models based on artificial intelligence and digital twins; III) digital technologies for climate change adaptation and mitigation. The results highlight a clear prevalence of applications aimed at climate adaptation, especially in the upstream stages of the supply chain, through the use of IoT sensors, open-source platforms, GIS systems, remote sensing, drones, and AI models for monitoring water stress, forecasting yields, managing plant health, and optimising irrigation. The findings on climate change mitigation primarily focus on improving input efficiency, reducing the carbon footprint, and valorising by-products within a circular bioeconomy framework.

The study highlights how digital technologies serve as a strategic lever to strengthen resilience, traceability, and competitiveness in the sector. At the same time, the findings underscore the need to develop more integrated approaches between adaptation and mitigation, as well as more advanced predictive and prescriptive models that systematically support decision-making processes and future climate reporting across the supply chain.

Keywords: climate change, digital technologies, extra virgin olive oil, olive production, systematic literature review

Paper type: Practical Paper

1 Introduction and Background

The extra virgin olive oil (EVOO) sector plays a strategic role in the European agri-food system, both nutritionally and economically, representing a key element of the Mediterranean cultural and landscape heritage (El Joumri et al., 2023). EVOO is widely recognised for its health benefits and high nutritional value. Within this framework, the European Union (EU) plays a central role in the global olive oil sector, accounting for approximately 65–70% of total production and generating an annual market value of approximately €7–8 billion, with EVOO being the segment with the highest added value (OIV, 2023). The EU also remains the world's leading exporter, with over 800,000 tonnes annually, primarily destined for the United States, China, and Japan. The sector also plays an important social role, involving more than 2.5 million producers and contributing to rural development and employment protection (Özdel et al., 2025).

However, the EVOO supply chain is increasingly exposed to the effects of climate change, including rising temperatures, prolonged droughts, greater climate variability, and extreme events, which directly impact olive tree physiology, yield stability, and oil quality (Kaniewski et al., 2023). Specifically, these phenomena lead to water stress, alterations in phenological phases, and the spread of plant diseases and pests, compromising both the quantity produced and the product's nutritional and sensory properties (Demissie et al., 2026). Furthermore, the high economic value of EVOO makes it vulnerable to fraud, which undermines consumer trust and producers' competitiveness (Parra-López et al., 2025).

In this context, digital technologies are emerging as strategic tools for developing resilient and efficient agri-food systems. The literature highlights the growing adoption of digital technologies across the supply chain, particularly in the upstream stages of cultivation and harvesting (Stillitano et al., 2019). Technologies such as remote sensing, Internet of Things (IoT)-based systems, artificial intelligence (AI), machine learning (ML), and decision support systems (DSS) enable monitoring of climate variables, assessing crop health, optimising irrigation, and supporting data-driven management (Lozano-Castellón et al., 2022), improving olive growers' adaptive capacity (Casado-Vara et al., 2018). Furthermore, tools such as blockchain enable traceability and authenticity throughout the supply chain, strengthening transparency and trust in the market (Salzano et al.,

2026; Parra-López et al., 2025). Despite this, the literature remains fragmented, with studies focusing on individual technologies or supply chain stages and limited integration between climate adaptation and mitigation objectives, underscoring the need for a systematic, structured analysis of the role of digital technologies in the EVOO sector (Violino et al., 2019).

In light of the identified gaps, this study proposes to analyse the role of digital technologies across the EVOO supply chain, with particular reference to their contributions to climate change mitigation and adaptation. From this perspective, the analysis, through a systematic literature review, examines how these technologies support resource efficiency and process improvement throughout the product life cycle, identifying the main research lines, existing gaps, and future research perspectives.

The study is structured as follows: Section 2 describes the methodology; Section 3 presents the results; and Section 4 presents the discussion and conclusions.

2 Methodology

The study, following the guidelines proposed by Tranfield et al. (2003), adopts a systematic literature review (SLR) methodology to analyse existing literature in a structured manner. The literature search was conducted on the Scopus database in January 2026, and studies were analysed by both author and conceptual content to extract relevant information (Webster and Watson, 2002).

The systematic literature review was guided by the following research questions:

- *(RQ1) What is the role of innovative digital technologies along the extra virgin olive oil supply chain?*
- *(RQ2) How do these technologies contribute to climate change mitigation and adaptation processes in terms of resource optimisation, process efficiency, and reduced environmental impacts?*

To answer the research questions, the study extracts, synthesises, and critically analyses the existing literature.

The Scopus database search was conducted using the following search term: (olive OR "olive oil sector" OR "edible oil" AND "climate change" OR "climate adaptation" OR "climate mitigation" AND "digital technologies" OR "artificial intelligence" OR "AI" OR "blockchain technology" OR Blockchain OR "Internet of Things" OR IoT). The following filters were applied to the database used: (i) document type: scientific articles, literature reviews, and conference proceedings; (ii) language: English; and (iii) full-text availability. The authors decided not to impose strict time constraints to capture the evolution of the literature on the topic.

The initial search identified 42 contributions. All records were screened by title and abstract. During this phase, 20 articles were excluded as they did not align with the research objectives. Therefore, 22 studies were eligible for full-text review.

During the eligibility phase, the full texts were analysed in detail, and two contributions were excluded because they did not fit the study's relevant disciplinary fields (Environmental Science, Computer Science, Business, Management and Accounting, Social Sciences, Economics, Econometrics, and Finance). After applying these criteria, the final literature corpus consisted of 20 articles included in the systematic review. The screening and selection process was conducted by two researchers to ensure the reliability of the review. In the event of disagreement, decisions were discussed until consensus was reached.

The full texts of the final literature were analysed to gather information on the scope, study approach, methodology, results, and suggestions for future research. This information was then organised into an Excel file.

In light of the research questions underlying the systematic literature review, the selected studies were examined using content analysis. The aim was to describe the bibliometric characteristics of the existing literature, identify the main areas of application, and analyse, for each area, the findings and future research perspectives.

The authors highlighted the bibliometric characteristics of the reviewed articles and subsequently presented the results of the three identified macro-thematic areas. Specifically, the analysis distinguished between: (i) digitalisation and new technologies for operational optimisation in olive growing; (ii) AI-based predictive models and Digital Twins for crop management; and (iii) climate change adaptation or mitigation.

3 Results

3.1 Bibliometric characteristics of the reviewed articles

The analysis shows that the publication years of the selected articles range from 2020 to 2026. Bibliometric analysis shows that between 2023 and 2025, there was a significant increase in publications on the role of

innovative digital technologies in climate change mitigation and adaptation processes along the extra-virgin olive oil supply chain, demonstrating that the topic has become more relevant to the scientific community. Specifically, one article was selected for 2020, while the authors analysed two studies for 2021. Four articles were selected for 2023, and five for 2024. 2025 was the year with the most selected contributions (eight). Finally, only one article was selected for 2026, as the literature review was conducted in January 2026, limiting the availability of contributions published that year. Regarding the typology, 14 are journal articles, 4 are conference proceedings, and 2 are reviews.

The methodologies adopted in the reviewed studies include 11 case studies, 5 conceptual analyses, and 2 systematic literature reviews.

For the empirical studies, the geographical contexts of application are distributed as follows: Italy (7 studies), Morocco (2), Portugal (2), Spain (1), Greece (1), Turkey (1), Cyprus (1), Algeria (1), Djibouti (1), and Nigeria (1).

3.2 Digitalisation and new technologies for operational optimisation in olive growing

The analysis highlights how integrating Agriculture 4.0 technologies in the olive sector has radically transformed the management of agronomic activities, promoting data-driven operational efficiency (Parra-López et al., 2025). Seven reviewed articles focus on operational optimisation in olive growing through the use of digital technologies. The analysis highlighted that the most commonly used technologies include smart sensors, IoT platforms, drones/UAVs, GIS systems, and remote sensing tools.

The review emphasises that the implementation of smart sensor networks forms the backbone of remote, continuous monitoring of olive grove conditions (Tzerakis et al., 2023). In particular, the effectiveness of LoRaWAN (Long Range Wide Area Network), a long-range, low-power communication protocol, makes it well-suited for large olive groves in remote areas where traditional network connectivity is limited (Varandas et al., 2020). These low-cost sensor networks enable real-time monitoring of vital parameters, including soil moisture, air temperature, and plant water status. In this sense, the use of internal and external sensors enables the assessment of plant temperature differences, providing crucial data to determine the actual need for pesticide treatments and to reduce chemical waste (Tzerakis et al., 2023; Varandas et al., 2020). The system's energy autonomy is often ensured by integrating photovoltaic panels, making the sensing nodes completely independent of the electricity grid and less polluting. The collected data is subsequently fed into open-source IoT platforms (such as ThingsBoard or Node-RED), which, through intuitive dashboards, inform farmers about irrigation needs and the risk of pest infestation, such as the olive fruit fly (Bactrocera oleae) (Tzerakis et al., 2023; Ammouri et al., 2019).

Among new technologies, the introduction of robotic systems and unmanned aerial vehicles (UAVs) is particularly important, as they aim to automate labour-intensive tasks and reduce labour costs. Unmanned aerial vehicles (UAVs) play a dual role: in addition to providing high-resolution mapping to identify water stress and nutrient deficiencies, they are used as autonomous sprayers. The latter allows for precise application of pesticides, drastically reducing water consumption, soil compaction, and airborne product drift (Demissie et al., 2026; Parra-López et al., 2025).

The use of precision agriculture, particularly the combined application of geographic information systems (GIS) and satellite remote sensing, enables site-specific management that optimises the use of production inputs such as water, fertilisers, and pesticides (Demissie et al., 2026; Granado-Díaz et al., 2024). Multispectral image analysis from satellites such as Sentinel-2, Landsat, and MODIS allows for large-scale tree health monitoring, facilitating the classification of phenological stages and accurate yield estimation (Demissie et al., 2026). GIS platforms support site-suitability assessments and the planning of optimal irrigation strategies, thereby improving the resilience of olive groves to climate change (Parra-López et al., 2025; Tzerakis et al., 2023). This data-driven approach ensures that each agronomic intervention is tailored to the actual needs of each specific plot of land, reducing environmental impact and preserving the quality of the final product (Parra-López et al., 2025; Demissie et al., 2026; Granado-Díaz et al., 2024).

3.3 AI-based predictive models and digital twins for crop management

Six reviewed articles analyse AI-based predictive models and digital twins for crop management, identifying relevant applications in yield estimation, water status monitoring, irrigation planning, and early detection of stress, disease, and plant diseases.

The application of Artificial Intelligence (AI) and advanced digital models in the olive sector is driving a paradigm shift toward data-driven management, enabling accurate forecasts and optimal resource coordination. These technologies offer sophisticated analytical tools for processing large volumes of agricultural

data, supporting olive growers in making strategic decisions such as yield estimation, disease detection, and quality control (Salzano et al., 2026; Preite et al., 2025).

Indeed, integrating AI techniques with remote sensing (RS) overcomes the limitations of traditional methods, improving the accuracy and scalability of growth parameter and crop yield estimation (Demissie et al., 2026). Machine learning algorithms, such as Random Forest (RF), are used to analyse vegetation indices derived from satellite or unmanned aerial vehicle (UAV) imagery to predict final yield. In particular, the use of convolutional neural networks (CNNs), such as the Mask R-CNN model applied to UAV imagery, enables tree canopy segmentation and biovolume estimation, facilitating accurate yield prediction at the individual plant level (Demissie et al., 2026; Parra-López et al., 2025). Similarly, AI supports the classification of phenological stages (based on the BBCH scale), achieving high reliability (F1 Scores up to 0.89 for olive trees), a key element in planning timely agronomic interventions. Furthermore, AI-based species distribution models predict significant shifts of suitable cultivation areas northward or to higher altitudes due to climate change-induced drought stress (Salzano et al., 2026; Özdel et al., 2025).

The creation of digital twins of the soil-plant-environment system allows modelling the complex flows and interactions between living and non-living entities. The integration of IoT sensors with neural networks, specifically Multi-Layer Perceptron (MLP), has demonstrated accuracy exceeding 97% in predicting soil water content (Preite et al., 2025). These models process continuous data on leaf temperature, stem size, and ambient humidity to provide data-driven irrigation recommendations that optimise water use and mitigate water stress. Advanced approaches that combine particle swarm optimisation (PSO) and neural networks (ANN) allow the system to dynamically learn from new datasets, automating irrigation and liquid fertilisation decisions without human intervention (Preite et al., 2025; Tawfeek et al., 2022).

Finally, the use of deep learning models has revolutionised plant health monitoring, enabling early detection of biotic and abiotic stresses. Convolutional neural networks (CNNs) are valuable tools for identifying and classifying images captured by IoT devices or drones, thereby facilitating automated diagnosis of leaf diseases (Salzano et al., 2026; Parra-López et al., 2025). The integration of remote sensing, AI, and UAVs is particularly effective for early, even pre-visual, detection of devastating pathogens such as Xylella fastidiosa and for monitoring traps for the olive fruit fly (Bactrocera oleae). Open-source IoT platforms can also generate automatic email alerts when climatic conditions (temperature between 15°C and 26°C and relative humidity >65%) favour the development of pest populations, allowing farmers to adopt targeted protective measures and reduce the indiscriminate use of pesticides (Salzano et al., 2026; Parra-López et al., 2025).

3.4 Adaptation or mitigation of climate change

The authors identified seven articles that analysed approaches to assessing climate change adaptation and mitigation using digital technologies.

The review results reveal a clear predominance of adaptation strategies over mitigation approaches. This trend reflects the urgency of responding to the immediate impacts of climate variability in the Mediterranean basin, where digital technologies are being used as primary tools to ensure resilience and yield stability under increasing environmental stress (Demissie et al., 2026).

Regarding adaptation, studies primarily focus on enhancing olive orchard resilience through monitoring and forecasting abiotic and biotic stresses. The integration of Artificial Intelligence (AI) and IoT sensors enables modelling of flows between living and non-living entities, predicting soil water content with greater than 97% accuracy, and providing irrigation recommendations crucial for drought mitigation (Preite et al., 2025). Advanced predictive models based on machine learning also highlight that climate change is causing northward shifts in areas suitable for cultivation and upward shifts to higher altitudes, influencing long-term strategic decisions in the planning of new plantations (Majikumna et al., 2024). In crop protection, digitalisation enables early warning systems for emerging adversities: algorithms based on climate rules send automatic notifications when conditions favour the proliferation of the olive fruit fly (Bactrocera oleae), while the use of mobile robots and drones facilitates the early identification of devastating pathogens such as Xylella fastidiosa (Tzerakis et al., 2023).

Conversely, the effects of climate change mitigation are addressed more indirectly and documented less quantitatively. Reducing greenhouse gas (GHG) emissions primarily involves optimising resource use (input efficiency): precise management of water, fertilisers, and pesticides, enabled by precision agriculture, reduces the overall carbon footprint of the production phase (Demissie et al., 2026). Furthermore, the literature identifies the potential of olive ecosystems to sequester carbon in biomass and soil, suggesting that advanced digital management could improve CO2 storage capacity (Majikumna et al., 2024; Granado-Díaz et al., 2024). A promising path to future sustainability is the circular bioeconomy, where digitalisation supports the valorisation

of by-products (olive oil mill wastewater, olive pomace, and pruning residues) for producing biofuels or bioactive compounds, thereby transforming residues into resources and reducing the environmental impact of the transformation phase. Although quantitative evidence remains limited, the integration of Industry 4.0 technologies with life cycle assessment (LCA) frameworks emerges as an enabler of future supply chain climate reporting (Parra-López et al., 2025).

4 Discussions and Conclusions

The study highlights an approach to integrating digital technologies into the EVOO supply chain, one that plays a central role, especially in the upstream phases, by supporting olive grove monitoring, crop management, yield forecasting, and the optimisation of agronomic practices. Specifically, the integration of IoT sensors, AI, remote sensing, UAVs, cloud platforms, and GIS systems enables more data-driven production management, improving operators' ability to promptly detect water stress, plant diseases, and changes in plant physiological conditions (Salzano et al., 2026; Demissie et al., 2026; Granado-Díaz et al., 2024).

The results show that the technologies adopted in the sector are primarily geared toward supporting climate change adaptation, strengthening the resilience of olive groves, and improving the efficient use of critical resources such as water, fertilisers, and pesticides. Specifically, the use of internal and external sensors, AI-based predictive models, and digital twins enables more accurate assessments of plant water and physiological status, thereby enabling more targeted and timely agronomic interventions (Preite et al., 2025; Tawfeek et al., 2022). At the same time, the study confirms a growing interest in more integrated climate mitigation approaches based on resource efficiency and the principles of the circular economy (Majikumna et al., 2024). In this context, digital technologies play an enabling role in the development of the bioeconomy, promoting the valorisation of olive mill byproducts, such as wastewater and pomace, for the extraction of high-value compounds or for the production of biofuels. This approach allows waste to be transformed into resources, thereby reducing the carbon footprint and environmental impacts associated with the transformation phase (Parra-López et al., 2025).

Furthermore, the findings highlight the opportunity to develop more structured digital governance systems, based on smart contracts and open-source cloud platforms, capable of connecting traditional agronomic practices with advanced technological tools and promoting more equitable access to innovation, even for small-scale producers (Salzano et al., 2026; Granado-Díaz et al., 2024). The creation of an integrated digital ecosystem, in which blockchain ensures data integrity, and AI provides high-precision predictive models (achieving accuracy levels above 97% for critical variables such as soil moisture), defines a new paradigm of "Traceability 4.0" (Parra-López et al., 2025). This integration not only protects EVOO against fraud and adulteration but also consolidates consumer trust and ensures the global competitiveness of Mediterranean olive cultivation, making it a resilient pillar in the face of pressing environmental and market challenges (Salzano et al., 2026; Demissie et al., 2026). Among the main managerial implications is the ability to forecast crop yields at the beginning of the year with an error margin of less than 20%, enabling advance economic planning and targeted crop management (Parra-López et al., 2025). Furthermore, the use of IoT platforms and AI models enables irrigation optimisation with an accuracy of over 97%, provides early warnings of pests such as the olive fruit fly, and drastically reduces pesticide use and operating costs (Preite et al., 2025; Tzerakis et al., 2023).

This systematic literature review has several limitations. First, the investigation was limited exclusively to the Scopus database, a choice motivated by the need to ensure full reproducibility and transparency of the search process. Although this approach ensures methodological rigour and consistency in the analysis, it may have excluded relevant contributions indexed in other databases. Furthermore, the decision to include only articles, reviews, and conference papers published in English may have further limited the inclusion of studies written in other languages. Consequently, the overall reconstruction of the digitalisation phenomenon in the EVOO sector may be only partially representative in terms of international coverage (Macke et al., 2019).

In light of the emerging evidence, one of the main directions for future research concerns the progressive shift towards increasingly predictive and prescriptive management models. Indeed, the synergistic integration of multi-scale data from heterogeneous platforms, such as satellite remote sensing, drones, and sensors, appears particularly promising for more accurately monitoring the physiological state of plants and optimising resource use (Demissie et al., 2026).

Furthermore, it will be crucial to develop phenology-aware AI models that incorporate the specific phenological phases of the olive tree to improve yield estimates and more effectively calibrate the use of production inputs (Demissie et al., 2026). Further expected developments include advances in automation and intelligent robotics, through the design of systems that support selective harvesting based on olive ripeness and

automated pruning. In this context, the use of autonomous sprayers and precision endotherapy techniques could also help combat vascular diseases, thereby significantly reducing chemical use (Parra-López et al., 2025).

These evolutionary trajectories outline relevant perspectives for the construction of an olive-growing sector more resilient to climate change, in which technology is configured not only as a lever for production efficiency, but also as a strategic factor for the competitiveness of the sector and the safeguarding of the cultural heritage connected to it (Granado-Díaz et al., 2024).

5 Ethics declaration

Ethical clearance was not required for this study, as it is based exclusively on a systematic literature review of publicly available sources.

6 AI declaration

The authors declare that no artificial intelligence tools were used in the analysis, processing and writing of this manuscript.

References

Ait Ahmed, N., Benmessaoud, Y., Hamdail, B.R. and Ezziyyani, M. (2024) "Intelligent System for Forecasting Climate Deterioration and Assessing the Potential Impact of Pollutants", *Medicon Agriculture & Environmental Sciences*, Vol. 6, No. 4, pp 40–51.

Alves, A., Morais, A.J., Filipe, V. and Pereira, J.A. (2021) "Intelligent monitoring and management platform for the prevention of olive pests and diseases, including IoT with sensing, georeferencing and image acquisition capabilities through computer vision".

Ammouri, J., Minet, P., Boudiaf, M., Bouzefrane, S. and Yacoub, M. (2019) "Self-Organizing Maps Applied to Soil Conservation in Mediterranean Olive Groves", in *PEMWN2019 - 8th IFIP/IEEE International Conference on Performance Evaluation and Modeling in Wired and Wireless Networks*.

Casado-Vara, R., Prieto, J., De la Prieta, F. and Corchado, J.M. (2018) "How blockchain improves the supply chain: Case study alimentary supply chain", *Procedia Computer Science*, Vol. 134, pp 393–398.

Demissie, W.A., Sebastiani, L. and Rossetto, R. (2026) "Integration of artificial intelligence and remote sensing for crop yield prediction and crop growth parameter estimation in Mediterranean agroecosystems: Methodologies, emerging technologies, research gaps, and future directions", *European Journal of Agronomy*, Vol. 173, 127894.

El Idrissi, A., Haidine, A., Dahbi, A. and Aqqal, A. (2022) "Smart Oasis: Using ICT Advanced Solutions for Societal and Economic Reconnection", *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. XLVIII-4/W3-2022, pp 47–53.

El Joumri, L., Labjar, N., Dalimi, M., Harti, S., Dhiba, D., El Messaoudi, N., ... and El Hajjaji, S. (2023) "Life cycle assessment (LCA) in the olive oil value chain: A descriptive review", *Environmental Development*, Vol. 45, 100800.

Granado-Díaz, R., Colombo, S., Romero-Varo, M. and Villanueva, A.J. (2024) "Farmers' attitudes toward the use of digital technologies in the context of agri-environmental policies", *Agricultural Systems*, Vol. 221, 104129.

Houlali, I., Gouskir, M., Elbir, M., Amhoud, A., Moubarik, A., Bouikhalene, B., Chaair, H., Rabi, S., Abdelaoui, O., Echajia, M., Bouhdadi, R. and Mbarki, M. (2014) "Effects of sunlight, temperature and adulteration on the quality of the virgin olive oil", *Journal of Materials and Environmental Science*, Vol. 5, No. 4, pp 1021–1026.

International Organisation of Vine and Wine (2023) "Annual Assessment of the World Vine and Wine Sector in 2023", International Organisation of Vine and Wine.

Kaniewski, D., Marriner, N., Morhange, C., Khater, C., Terral, J.F., Besnard, G., ... and Cheddadi, R. (2023) "Climate change threatens olive oil production in the Levant", *Nature Plants*, Vol. 9, No. 2, pp 219–227.

Kleyböcker, A. and Vamvakeridou-Lyroudia, L. (2025) "Editorial: Circular economy in the context of water smart industrial symbioses (project ULTIMATE)", *Water Science & Technology*, Vol. 92, No. 10, pp iii–v.

Lozano-Castellón, J., López-Yerena, A., Domínguez-López, I., Siscart-Serra, A., Fraga, N., Sámano, S., ... and Pérez, M. (2022) "Extra virgin olive oil: A comprehensive review of efforts to ensure its authenticity, traceability, and safety", *Comprehensive Reviews in Food Science and Food Safety*, Vol. 21, No. 3, pp 2639–2664.

Macke, J. and Genari, D. (2019) "Systematic literature review on sustainable human resource management", *Journal of Cleaner Production*, Vol. 208, pp 806–815.

Majikumna, K.U., Zineddine, M. and Alaoui, A.E.H. (2024) "Olive tree drought stress: A systematic review", *Journal of Water and Climate Change*, Vol. 15, No. 12, pp 5741–5762.

Mennella, V.G.G., Pacicco, L. and Grohmann, D. "Evoluzione degli ecosistemi naturali e agricoli in relazione ai cambiamenti climatici", *Associazione Italiana di Ingegneria Agraria*.

Özdel, M.M., Ustaoğlu, B. and Cürebal, İ. (2025) "Climate-Driven Futures of Olive (Olea europaea L.): Machine Learning-Based Ensemble Species Distribution Modelling of Northward Shifts Under Aridity Stress", *Plants*, Vol. 14, 3774.

Paré, G. and Kitsiou, S. (2017) "Methods for literature reviews", in *Handbook of eHealth Evaluation: An Evidence-Based Approach* [Internet], University of Victoria.

Parra-López, C., Ben Abdallah, S., Hassoun, A., Jagtap, S., Garcia-Garcia, G., Ben Hassen, T., Trollman, H., Trollman, F. and Carmona-Torres, C. (2025) "Digital and Industry 4.0 technologies in olive farming and industry: Recent applications and future outlook", *Smart Agricultural Technology*, Vol. 12, 101376.

Payab, A.H. and Türker, U. (2025) "Resilience of rainfed agriculture in northern part of Cyprus: a comprehensive drought severity level impact analysis and adaptive strategies", *Environment*, *Development and Sustainability*.

Preite, L., Oddi, G., Galaverni, M., Belli, L., Rodolfi, M., Davoli, L., Lino, T., Vignali, G., Ferrari, G. and Ganino, T. (2025) "An IoT Data-Driven Framework for Olive Groves Monitoring and Irrigation Management", *IFAC PapersOnLine*, Vol. 59, No. 23, pp 279–284.

Salzano, F., Bonavita, N., Di Florio, C., Esposito, V., Guglielmi, E., Oliveto, R., Pareschi, R. and Scalabrino, S. (2026) "Nova Tellus: Integrating Blockchain and AI for Advanced Agricultural Management", *SN Computer Science*, Vol. 7, 108.

Stillitano, T., Falcone, G., De Luca, A.I., Piga, A., Conte, P., Strano, A. and Gulisano, G. (2019) "A life cycle perspective to assess the environmental and economic impacts of innovative technologies in extra virgin olive oil extraction", *Foods*, Vol. 8, No. 6, 209.

Tawfeek, M.A., Alanazi, S. and Abd El-Aziz, A.A. (2022) "Smart Greenhouse Based on ANN and IOT", *Processes*, Vol. 10, No. 11, 2402.

Tranfield, D., Denyer, D. and Smart, P. (2003) "Towards a methodology for developing evidence-informed management knowledge by means of systematic review", *British Journal of Management*, Vol. 14, No. 3, pp 207–222.

Tzerakis, K., Psarras, G. and Kourgialas, N.N. (2023) "Developing an Open-Source IoT Platform for Optimal Irrigation Scheduling and Decision-Making: Implementation at Olive Grove Parcels", *Water*, Vol. 15, No. 9, 1739.

Varandas, L., Faria, J., Gaspar, P.D. and Aguiar, M.L. (2020) "Low-Cost IoT Remote Sensor Mesh for Large-Scale Orchard Monitorization", *Journal of Sensor and Actuator Networks*, Vol. 9, No. 3, 44.

Violino, S., Pallottino, F., Sperandio, G., Figorilli, S., Antonucci, F., Ioannoni, V., ... and Costa, C. (2019) "Are the innovative electronic labels for extra virgin olive oil sustainable, traceable, and accepted by consumers?", *Foods*, Vol. 8, No. 11, 529.

Webster, J. and Watson, R.T. (2002) "Analyzing the past to prepare for the future: Writing a literature review", *MIS Quarterly*, Vol. 26, No. 2, pp xiii–xxiii.

Yunusa, S.U. and Wakili, B.S. (2023) "Development of lignocellulosic-plastic composite from rice husk and polyethylene", *Cleaner and Circular Bioeconomy*, Vol. 6, 100054.

Mapping Circular Employability Skills in Agri-Food Value Chains from Large-Scale Labour-Market Data

Francesco Smaldone[1], Stefania Supino[2], Daniela Sica[2], Benedetta Esposito[2]
[1] G. Marconi University of Rome, Rome, Italy
[2] San Raffaele University Rome, Rome, Italy

Abstract

The twin transition, combining circular economy and digital transformation, is reshaping the food industry and generating new skill requirements. While previous studies have examined circular and digital technologies in agri-food systems, limited attention has been paid to the employability skills needed to support their implementation. This study addresses this gap by profiling the skills required for the twin transition in the food processing phase of the agri-food supply chain. The research adopts a big-data-driven approach based on 6,473 job advertisements collected from major online recruitment platforms. Starting from ESCO occupations, relevant food processing job profiles were selected and integrated with keyword-based searches. The textual corpus was analysed through natural language processing, BERTopic, hierarchical clustering, machine learning and SHAP analysis. Skills were classified into three semantic topics - hard, soft and transversal skills - and three cognitive clusters: industry-specific, circular and digital competences. Findings show that labour demand is dominated by hard and industry-specific skills, especially production, quality management, safety standards and packaging. Circular competences are increasingly present, whereas digital skills remain comparatively underrepresented, despite their strategic role in traceability, monitoring and efficiency. Advanced traceability, LCA-guided eco-design and industrial symbiosis emerge as high-impact but still underdeveloped competence areas. The study contributes to the literature by framing employability skills as strategic assets for enabling the twin transition. It also provides practical implications for firms, policymakers and education providers, highlighting the need to invest in targeted skill portfolios able to translate circular and digital innovation into sustainable performance.

Keywords: twin transition; employability skills; circular economy; digital technologies; food industry

Paper type: Academic Research Paper

1 Introduction

The global economy is undergoing a profound transformation driven by the twin transition (TT), understood as the simultaneous and interconnected green and digital transition aimed at reducing companies' carbon footprint. Rather than representing a merely technological shift, TT constitutes a systemic transformation that affects industrial structures, governance mechanisms, and labour markets (Trevisan et al., 2024). It combines two complementary trajectories: digitalisation, based on data-driven technologies, automation, and artificial intelligence, and circularity, aimed at decoupling economic growth from resource depletion through reuse, recycling, and regenerative business models (Uhrenholt et al., 2022).

At the European level, these transitions are increasingly viewed as key levers for competitiveness, sustainability, and resilience (European Commission, 2020; Vitti et al., 2025). The European Green Deal and the Circular Economy Action Plan promote the path towards climate neutrality by 2050, while the European Skills Agenda stresses the need for upskilling and reskilling to prepare the workforce for the new demands generated by digitalisation and sustainability (Cavallini & Soldi, 2023).

Academic research has also paid growing attention to the relationship between circular economy (CE) and digital technologies (DT). Studies show that DT can support circularity by improving traceability, resource efficiency, and circular business models (Kristoffersen et al., 2020; Tabares et al., 2025). However, the integration between digital and circular transitions is still hindered by several barriers, including technological immaturity, financial constraints, fragmented governance, weak policy coordination, and, crucially, the lack of adequate skills (Sica et al., 2024; Trevisan et al., 2023).

The shortage of skilled professionals is therefore a central obstacle to the TT. Scholars emphasise that this transition requires not only advanced technical expertise, but also a combination of digital, circular, and transversal skills (Acerbi et al., 2022; Beducci et al., 2024; Trevisan et al., 2024). Skills represent the human dimension of transformation, enabling organisations to convert sustainability and digital challenges into opportunities for innovation and competitiveness (Trevisan et al., 2025). Existing studies have begun to classify and map the skills needed for circular and digital transitions, proposing different frameworks (Trevisan et al.,

2024; Beducci et al., 2024; Vitti et al., 2025). Nevertheless, much of this literature focuses on education, higher education, and vocational training, thus privileging the supply side of skills development.

By contrast, less attention has been devoted to the demand side, namely the skills actually required by employers and signalled by the labour market. This perspective is captured by the concept of employability skills (ES), which refers to the set of knowledge, competencies, attitudes, and experiences that allow individuals to obtain and maintain employment while contributing effectively to organisations (Fugate et al., 2004; Pool & Sewell, 2007; Smaldone et al., 2022). Unlike narrowly defined technical qualifications, ES include hard, soft, and transversal dimensions, offering a more comprehensive lens through which to interpret the labour market implications of TT (Smaldone et al., 2025).

Despite their relevance, ES remain underexplored in relation to both circular and digital transitions. Existing studies have examined ES either in circular contexts (Smaldone et al., 2025) or in digital contexts (Smaldone et al., 2022), but no research has systematically analysed their combined implications for the TT. This gap is particularly significant because the mismatch between educational provision and labour market demand may reinforce skill shortages and reduce firms' capacity to implement sustainable and digital innovations (Trevisan et al., 2025).

A further limitation concerns the sectoral focus of existing research. Most studies on TT-related skills have concentrated on manufacturing, examining competences linked to remanufacturing, recycling, design for circularity, and digital integration (Acerbi et al., 2022; Beducci et al., 2024; Trevisan et al., 2024). The food industry has received far less attention, despite being one of the most environmentally impactful sectors and, at the same time, one of the most promising contexts for circular and digital strategies. The sector contributes substantially to greenhouse gas emissions, waste generation, and resource inefficiency, but it also offers opportunities for by-product valorisation, blockchain-based traceability, eco-design, nutrient cycling, and digital platforms for waste reduction (Esposito et al., 2020; Scandurra et al., 2023; Myshko et al., 2024; Farace & Tarabella, 2024; Sica et al., 2024).

In the transformation phase of the agri-food supply chain, the TT can generate particularly relevant benefits. Digitalisation can improve monitoring, transparency, and efficiency, while circular approaches can reduce environmental pressures through resource recovery and more sustainable production models. However, existing studies on CE and DT in the agri-food sector have mainly examined technological implementation, without adequately considering the role of skills in enabling the transition (Guidani et al., 2024; Padthar et al., 2024; Esposito et al., 2025). As a result, no prior research has systematically profiled the ES required to support TT in the food industry.

Against this background, the present study addresses this gap by profiling the employability skills required for the TT in the food industry, with specific attention to the transformation phase of the agri-food supply chain. Adopting a big-data-driven methodology, the study analyses thousands of job advertisements collected from major recruitment platforms and applies text mining and topic modelling techniques to identify relevant ES. These skills are then classified into semantic topics, including hard, soft, and transversal skills, and into cognitive clusters, namely industry-specific, circular, and digital skills. In doing so, the study develops a comprehensive taxonomy of the ES required by the European labour market to support the circular and digital transformation of the food industry.

2 Research Methodology

The methodological design was developed to ensure both representativeness and alignment with an institutional framework. The analysis began with a systematic screening of occupations listed in the European Skills, Competences, Qualifications and Occupations (ESCO) database, following prior studies (Akyazi et al., 2023). The focus was restricted to the transformation phase of the food industry, defined as the industrial stage in which agricultural raw materials are converted into intermediate or final food products through processing, refinement, preservation, and packaging. This phase was selected because it represents a crucial interface between upstream production and downstream distribution, where CE practices and DT are particularly relevant. All ESCO occupations related to the agri-food sector were first identified, and only those directly connected to food transformation were retained. Occupations linked exclusively to upstream or downstream activities were excluded. To ensure reliability, two independent researchers coded an initial sample of 30 job profiles, and Krippendorff's alpha exceeded the accepted threshold of 0.8, confirming satisfactory intercoder agreement (Krippendorff, 1980). The remaining profiles were then screened by the researchers, while final inclusion required unanimous agreement. The selected occupations were also cross-checked on major recruitment platforms. Since ESCO may not fully capture rapidly evolving labour market needs (Akyazi et al., 2022, 2023), an additional keyword-based search was conducted using terms such as "food", "agri-food",

“agrifood”, “agrofood”, and related variants. Job advertisements were collected from leading online recruitment platforms between February and July 2025 through a custom R-based scraping protocol. Around 8,000 unique postings related to food processing were retrieved. After removing duplicates, incomplete entries, and advertisements lacking substantive descriptions, the final dataset included 6,473 job ads, enriched with metadata such as occupation title, country, date, sector, and stated requirements.

The textual corpus was then prepared through a multi-step natural language processing routine, including lowercasing, removal of punctuation and numerals, language detection, tokenisation, and lemmatisation. Food-industry-specific collocations were preserved to maintain domain accuracy (Smaldone et al., 2022). Semantic validation relied on a lexicon of 35 pre-identified skills relevant to food processing, covering technical, managerial, transversal, and soft categories. A stratified keyword-matching procedure was applied, and disagreements were resolved through consensus discussion to strengthen internal validity (Moore, 2017).

The validated dataset was analysed using advanced computational techniques. BERTopic, based on BERT embeddings, was used to identify latent semantic structures in the corpus (Devlin et al., 2019; Grootendorst, 2022). This produced three macro-topics consistent with employability literature: hard skills, soft skills, and transversal skills. In parallel, agglomerative hierarchical clustering with cosine distance generated three macro-clusters: industry-specific, circular, and digital. This dual structure provided complementary perspectives on the skill space. Further multivariate analyses, including Multiple Correspondence Analysis, Principal Component Analysis, and Linear Discriminant Analysis, were used to examine latent patterns and predictive strength. Chi-square tests confirmed a significant relationship between topics and clusters, while Gini and Theil indices highlighted inequalities in the distribution of skills.

Cognitive profiling was carried out through heatmaps, radar charts, and skill networks, revealing the semantic relationships and weighted interconnections among competences. Network analysis distinguished strong and weak ties, showing the centrality of digitalisation and circularity in shaping employability in food processing (Smaldone et al., 2025). The predictive stage applied supervised machine learning, treating each skill as an observation and using topics and clusters as prediction targets. Decision Trees, Random Forests, and Support Vector Machines were tested (Cortes & Vapnik, 1995), with SHAP values used to interpret model outputs (Lundberg & Lee, 2017). Models were trained on an 80/20 stratified split, and Random Forest classifiers showed the strongest robustness. SHAP analysis identified the most discriminative skills, including communication, quality management, energy efficiency, and traceability systems.

Finally, the prescriptive stage conceptualised skills as a competence portfolio to support managerial decision-making. The Opportunity Map combines impact and prevalence to distinguish strategic opportunities, established leverage skills, maintenance skills, and marginal competencies. ROI analysis estimates the expected return from increasing the presence of a skill, comparing predictive and distributional benefits with its current baseline prevalence. What-if and sensitivity analyses simulate the effects of increasing specific skills, showing potential gains, trade-offs, and leverage effects. Lastly, the Concentration Risk gauge uses the Herfindahl–Hirschman Index to assess whether reliance on a narrow set of competences creates systemic fragility, with lower post-portfolio values indicating a more diversified and resilient skill base.

3 Results and Discussions

Results provide a comprehensive overview of the employability skills landscape within the food processing phase, combining descriptive distributions, inequality measures, predictive modelling and prescriptive scenarios. Overall, the distribution of competences across topics and clusters appears relatively balanced, although some relevant asymmetries emerge. Hard skills dominate the landscape, with frequencies exceeding 3,000 postings at their peak, compared with approximately 2,000 postings for both soft and transversal skills. A similar pattern emerges at the cluster level: industry-specific competences record more than 3,000 mentions, whereas circular and digital competences stabilise closer to 2,000 mentions. However, no single topic or cluster exceeds 30% of the total distribution, suggesting a fragmented rather than uniform skill structure.

The strongest intersections occur between hard skills and industry-specific competences, with 13,653 occurrences, followed by transversal skills and circular competences, with 11,306 occurrences. This indicates that traditional technical competences still play a central role in food processing employability, particularly in relation to production, processing, quality, safety and packaging. At the same time, transversal skills appear especially relevant when circular practices are involved, suggesting that the circular transition requires not only specialised technical knowledge, but also broader capabilities such as adaptability, collaboration, problem-solving and systems thinking.

Hard skills account for nearly 40% of job advertisements, confirming their predominance in labour demand. Soft skills, including communication and teamwork, appear in lower but still significant proportions, while

transversal skills are less frequent but strategically important for managing change and innovation. From a cluster perspective, industry-specific and circular competences together represent more than 70% of total demand, whereas digital skills remain close to 20%. This result confirms a substantive digital gap in food processing and supports previous evidence identifying digital immaturity as a key bottleneck in agri-food systems (Sica et al., 2024). The findings also reflect a structural skills gap, whereby firms retain strong compliance and operational capabilities but invest less in the competences needed for innovation and resilience (Trevisan et al., 2023).

The inequality analysis further clarifies this picture. The Gini index reaches 0.15 and the Theil index is approximately 0.05, indicating a moderate but meaningful concentration of skills. A limited number of competences, particularly industrial processing, safety standards and quality management, account for a disproportionate share of labour market demand. Chi-square residuals confirm this imbalance: circular competences are significantly over-represented in the circular cluster, with a residual of +66.82 for transversal skills, whereas digital competences show under-representation, with a residual of –41.64 for soft skills. These disparities suggest that circular priorities are increasingly embedded in food processing job profiles, while digitalisation, contrary to initial expectations, remains less consolidated.

The analysis of skill relationships reinforces this interpretation. Central competences are still dominated by food packaging and safety standards, while emerging digital competences such as smart sensors and IoT remain peripheral, despite their recognised potential to improve transparency, traceability and efficiency (Pakseresht et al., 2023; Shrivastava et al., 2023; Zhao et al., 2025). This confirms a relevant paradox: although digital technologies offer significant opportunities for the twin transition, their practical adoption is constrained by limited absorptive capacity and insufficient human capital (Kumar et al., 2021; Annosi et al., 2020).

Predictive modelling provides further support for these results. Accuracy rates approach 75% for both topic and cluster classification, while precision is close to 80%, confirming the predictive validity of the skill dataset. Hard skills are classified with higher accuracy, with 13 correct predictions, whereas soft and transversal skills show greater overlap and are more frequently misclassified between Topics 1 and 2. This reflects both methodological complexity and labour market reality: industrial competences tend to be more codified, while transversal competences cut across different roles and occupational profiles.

The most influential skills for topic classification are communication skills, years of experience, production efficiency, quality management and traceability systems. By contrast, smart sensors, greenhouse gas emissions and ingredient traceability are the most decisive for cluster classification. These findings confirm that employability in food processing is multidimensional, combining technical, soft, transversal, circular and digital competences. They also support the view that hybrid profiles, able to integrate technical expertise with adaptability and digital literacy, are increasingly relevant for competitiveness in changing labour markets (Van Laar et al., 2017; Shakina et al., 2021; Xu et al., 2022).

The prescriptive analysis translates these findings into strategic indications for workforce development. Eco-design, industrial symbiosis and advanced traceability emerge as high-impact but underrepresented competences, offering substantial potential returns. The selected skill portfolio shows benefit–cost ratios above 1.05, indicating that investment in these competences could generate value greater than their current prevalence. Scenario analysis shows that a 20% increase in transversal and circular competences would increase their representation by almost 2% across topics and 1.5% across clusters, contributing to a more balanced skill ecosystem. Sensitivity analysis further confirms that modest increases of 10% in competences such as eco-design and ingredient traceability produce measurable effects, whereas further investment in already dominant hard skills generates negligible marginal returns.

Finally, the concentration risk assessment, based on the Herfindahl–Hirschman Index, places the current configuration at 0.130, indicating low but still observable concentration risk. After portfolio diversification, the index decreases to 0.031, showing a clear reduction in systemic fragility and a stronger stabilising effect produced by redistributing competences across categories. Overall, the results reveal a food processing sector that is more mature in circularity than in digitalisation. The decisive lever is not simply the adoption of more technologies, but the development of targeted skill bundles able to transform digital tools into circular performance.

4 Conclusions

This study profiled the employability skills required to support the TT in the food industry, focusing specifically on the transformation phase of the agri-food supply chain. Through big-data text mining and AI-driven modelling, the analysis examined the distribution of hard, soft, and transversal skills, while also assessing their predictive relevance, trade-offs, and role within broader competence portfolios. The findings clearly show

that employability skills are a decisive factor in enabling firms to manage the convergence between circularity and digitalisation. Without adequate and diversified skill sets, food firms risk remaining anchored to traditional industrial models and failing to exploit the full potential of the TT.

The study offers relevant theoretical and managerial contributions. From a theoretical perspective, it enriches the literature on circular and digital transitions by placing employability skills at the centre of the debate. Skills are not interpreted as fixed qualifications, but as dynamic and strategic assets that shape firms' absorptive capacity and their ability to transform sustainability and digital pressures into innovation and competitive advantage. From a managerial perspective, the findings suggest that human capital should be treated as a form of strategic infrastructure. A diversified skill portfolio can reduce concentration risks, strengthen organisational resilience, and generate higher returns when investments are directed towards underrepresented but high-impact competences.

The study also presents some limitations. First, although job advertisements provide valuable evidence of labour market demand, they mainly reflect employer signalling rather than the actual competencies possessed by workers. Future research should therefore triangulate these data with employee surveys, training outcomes, and longitudinal career paths. Second, the analysis focuses only on the transformation phase of the agri-food supply chain. Further studies could extend the investigation to upstream production and downstream distribution and consumption, enabling a more comprehensive understanding of skill complementarities across the whole chain. Third, despite the robustness of the methodological framework, the results remain influenced by the quality, language, and structure of online recruitment data, as well as by the rapid evolution of labour markets.

5 Ethics Declaration

No ethical clearance was required for this study, as it did not involve human participants, personal data, or sensitive information requiring ethical approval.

6 AI Declaration

The authors used DeepL only for language proofreading and stylistic refinement. The intellectual content, analysis, interpretation, and final responsibility for the manuscript remain entirely with the authors.

References

Acerbi, F., Hofmann Trevisan, A., Vitti, M., Caldarelli, P., Dukovska-Popovska, I., Downes, S., ... & Sassanelli, C. (2024, September). Market needs for a circular transition: implemented practices and required skills. In IFIP International Conference on Advances in Production Management Systems (pp. 262-274). Cham: Springer Nature Switzerland.

Acerbi, F., Rossi, M., & Terzi, S. (2022). Identifying and assessing the required I4. 0 skills for manufacturing companies' workforce. *Frontiers in Manufacturing Technology*, 2, 921445.

Audrin, B., Audrin, C., & Salamin, X. (2024). Digital skills at work–Conceptual development and empirical validation of a measurement scale. *Technological Forecasting and Social Change*, 202, 123279.

Bag, S., & Rahman, M. S. (2024). Navigating circular economy: Unleashing the potential of political and supply chain analytics skills among top supply chain executives for environmental orientation, regenerative supply chain practices, and supply chain viability. *Business Strategy and the Environment*, 33(2), 504-528.

Bassi, F., & Guidolin, M. (2021). Resource efficiency and Circular Economy in European SMEs: Investigating the role of green jobs and skills. *Sustainability*, 13(21), 12136.

Beducci, E., Acerbi, F., Pinzone, M., & Taisch, M. (2024). Unleashing the role of skills and job profiles in circular manufacturing. *Journal of Cleaner Production*, 449, 141456.

Ben Youssef, A., Dahmani, M., & Ragni, L. (2022). ICT use, digital skills and students' academic performance: Exploring the digital divide. *Information*, 13(3), 129.

Bodescu, D., Sîrghea, A., Raţu, R. N., Chiruţă, C., Mălăncuş, R. N., Donosă, D., & Robu, A. D. (2024). Relevant skills for employment and entrepreneurship in the agri-food sector. *Sustainability*, 16(2), 787.

Burger, M., Stavropoulos, S., Ramkumar, S., Dufourmont, J., & van Oort, F. (2019). The heterogeneous skill-base of circular economy employment. *Research Policy*, 48(1), 248-261.

Buyukyazici, D., & Quatraro, F. (2025). The skill requirements of the circular economy. Ecological *Economics*, 232, 108559.

Cavallini, S., & Soldi, R. (2023). Europe's Circular Economy and its Pact for Skills: Workin g Together for an Inclusive and Job-Rich Transition.

Cheng, C., Gao, Q., Ju, K., & Ma, Y. (2024). How digital skills affect farmers' agricultural entrepreneurship? An explanation from factor availability. *Journal of Innovation & Knowledge*, 9(2), 100477.

Christiaensen, L., Rutledge, Z., & Taylor, J. E. (2021). The future of work in agri-food. *Food Policy*, 99, 101963.

Clarke, M. (2018). Rethinking graduate employability: The role of capital, individual attributes and context. *Studies in higher education*, 43(11), 1923-1937.

Cordella, A., Gualdi, F., & van de Laar, M. (2024). Digital skills within the Public Sector: A missing link to achieve the Sustainable Development Goals (SDGs). *Information Polity*, 29(1), 13-33.

Cortes, C., & Vapnik, V. (1995). Support-vector networks. Machine learning, 20(3), 273-297.

De los Rios, I. C., & Charnley, F. J. (2017). Skills and capabilities for a sustainable and circular economy: The changing role of design. *Journal of Cleaner Production*, 160, 109-122.

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). Bert: Pre-training of deep bidirectional transformers for language understanding. In Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies, volume 1 (long and short papers) (pp. 4171-4186).

Đorđević, B., Milanović Zbiljić, S., & Radosavljević, M. (2025). Impact of the Digital Skills on Employability: Cross-Sectional Analysis. *Economies*, 13(7), 196.

Esposito, B., Malandrino, O., Sessa, M. R., & Sica, D. (2023). Big Data and Digital Technologies for Circular Economy in the Agri-food Sector. In Big Data and Decision-Making: Applications and Uses in the Public and Private Sector (pp. 121-141). Emerald Publishing Limited.

Esposito, B., Smaldone, F., Raimo, N., Sica, D., & Vitolla, F. (2025). Enhancing circular economy disclosure towards Net Zero: the role of circular economy capabilities. *International Journal of Productivity and Performance Management*.

Farace, B., & Tarabella, A. (2024). Exploring the role of digitalization as a driver for the adoption of circular economy principles in agrifood SMEs–an interpretive case study. *British Food Journal*, 126(1), 409-427.

Górecki, J. (2019). Simulation-Based Positioning of Circular Economy Manager's Skills in Construction Projects. *Symmetry*, 12(1), 50.

Grootendorst, M. (2022). BERTopic: Neural topic modeling with a class-based TF-IDF procedure. arXiv. arXiv preprint arXiv:2203.05794.

Guerreschi, A., Piras, L., & Heck, F. (2023). Barriers to efficient knowledge transfer for a holistic circular economy: insights towards green job developments and training for young professionals. *Youth*, 3(2), 553-578.

Guyot Phung, C. (2019). Implications of the circular economy and digital transition on skills and green jobs in the plastics industry. Field Actions Science Reports. The Journal of Field Actions, (Special Issue 19), 100-107.

Helsper, E. J., & Van Deursen, A. J. A. M. (2015). Digital skills in Europe: Research and policy. *Digital divides: The new challenges and opportunities of e-inclusion*, 195(125), 45-63.

Holmes, L. (2013). Competing perspectives on graduate employability: possession, position or process?. *Studies in higher education*, 38(4), 538-554.

Janssens, L., Kuppens, T., & Van Schoubroeck, S. (2021). Competences of the professional of the future in the circular economy: Evidence from the case of Limburg, Belgium. *Journal of Cleaner Production*, 281, 125365.

Krippendorff, K. (1980), "Validity in content analysis", in Mochmann, E. (Ed.), Computerstrategien fÃ¼r die kommunikationsanalyse, Campus, Frankfurt, pp. 69-112, available at: http://repository.upenn.edu/asc_papers/291.

Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. Advances in neural information processing systems, 30.

Maer-Matei, M. M., Zamfir, A. M., & Mocanu, C. (2024). Skills and Jobs in Demand for the Circular Economy. In Circular Economy and Sustainable Development: A Necessary Nexus for a Sustainable Future (pp. 59-71). Cham: Springer International Publishing.

Martin, H. (2025). Skills for a Circular Economy Transformation—A Regional Case Study From Sweden. *International Journal of Training and Development*.

Mayor, L., Lindner, L. F., Knöbl, C. F., Ramalho, A., Berruto, R., Sanna, F., ... & Busato, P. (2022). Skill needs for sustainable agri-food and forestry sectors (I): assessment through European and national focus groups. *Sustainability*, 14(15), 9607.

Merriam-Webster dictionary, 2024. Skill [WWW Document]. URL. https://www. merriam-webster.com/dictionary/skill.

Moore, D. S. (2017). Tests of chi-squared type. In Goodness-of-fit-techniques (pp. 63-96). Routledge.

Munzert, S., Rubba, C., Meißner, P., & Nyhuis, D. (2015). Automated data collection with R: A practical guide to web scraping and text mining (Vol. 1). West Sussex: Wiley.

Myshko, A., Checchinato, F., Colapinto, C., Finotto, V., & Mauracher, C. (2024). Towards the twin transition in the agri-food sector? Framing the current debate on sustainability and digitalisation. *Journal of Cleaner Production*, 452.

Ordóñez de Pablos, P., Anshari, M., & Almunawar, M. N. (Eds.). (2024). Harnessing Green and Circular Skills for Digital Transformation. IGI Global.

Padthar, S., Naruetharadhol, P., Srisathan, W. A., & Ketkaew, C. (2024). From linear to circular economy: Embracing digital innovations for sustainable agri-food waste management among farmers and retailers. *Resources*, 13(6), 79.

Pagani, L., Argentin, G., Gui, M., & Stanca, L. (2016). The impact of digital skills on educational outcomes: evidence from performance tests. *Educational studies*, 42(2), 137-162.

Perçin, S. (2022). Evaluating the circular economy–based big data analytics capabilities of circular agri-food supply chains: the context of Turkey. *Environmental Science and Pollution Research*, 29(55), 83220-83233.

Perçin, S. (2023). Identifying barriers to big data analytics adoption in circular agri-food supply chains: a case study in Turkey. *Environmental Science and Pollution Research*, 30(18), 52304-52320.

Pirzada, K., & Khan, F. (2013). Measuring relationship between digital skills and employability. *European Journal of Business and Management*, 5(24).

Ramalho Ribeiro, A., Goodburn, B., Mayor, L., Lindner, L. F., Knöbl, C. F., Trienekens, J., ... & Busato, P. (2023). Skill needs for sustainable agri-food and forestry sectors (II): Insights of a European survey. *Sustainability*, 15(5), 4115.

Rhoades, S. A. (1993). The herfindahl-hirschman index. Fed. Res. Bull., 79, 188.

Sassanelli, C., Acerbi, F., Pellucchi, I., Dimitrova, D., Downes, S., Dukovska-Popovska, I., ... & Terzi, S. (2024, July). A Survey Investigating Skills, Job Profiles, and Education in Circular Economy: A Practice Perspective. In IFIP International Conference on Product Lifecycle Management (pp. 56-70). Cham: Springer Nature Switzerland.

Scandurra, F., Salomone, R., Caeiro, S., & Gulotta, T. M. (2023). The maturity level of the agri-food sector in the circular economy domain: A systematic literature review. *Environmental Impact Assessment Review*, 100, 107079.

Shakina, E., Parshakov, P., & Alsufiev, A. (2021). Rethinking the corporate digital divide: The complementarity of technologies and the demand for digital skills. *Technological Forecasting and Social Change*, 162, 120405.

Sica, D., Esposito, B., Malandrino, O., & Supino, S. (2024). The role of digital technologies for the LCA empowerment towards circular economy goals: a scenario analysis for the agri-food system. *The International Journal of Life Cycle Assessment*, 29(8), 1486-1509.

Smaldone, F., Ippolito, A., Lagger, J., & Pellicano, M. (2022). Employability skills: Profiling data scientists in the digital labour market. *European Management Journal*, 40(5), 671-684.

Smaldone, F., Lagger, J., & Mainolfi, G. (2025). Profiling employability skills for life cycle professionals amid the circular and digital economy. *The International Journal of Life Cycle Assessment*, 1-18.

Stofkova, J., Poliakova, A., Stofkova, K. R., Malega, P., Krejnus, M., Binasova, V., & Daneshjo, N. (2022). Digital skills as a significant factor of human resources development. *Sustainability*, 14(20), 13117.

Suarta, I. M., Suwintana, I. K., Sudiadnyani, I. G. A. O., & Sintadevi, N. P. R. (2024). Employability and digital technology: what skills employers want from accounting workers?. *Accounting Education*, 33(3), 274-295.

Summerton, L., Clark, J. H., Hurst, G. A., Ball, P. D., Rylott, E. L., Carslaw, N., ... & McElroy, C. R. (2019). Industry-informed workshops to develop graduate skill sets in the circular economy using systems thinking. *Journal of Chemical Education*, 96(12), 2959-2967.

Sumter, D., de Koning, J., Bakker, C., & Balkenende, R. (2021). Key competencies for design in a circular economy: Exploring gaps in design knowledge and skills for a circular economy. *Sustainability*, 13(2), 776.

Tabares, S., Parida, V., & Chirumalla, K. (2025). Twin transition in industrial organizations: Conceptualization, implementation framework, and research agenda. *Technological Forecasting and Social Change*, 213, 123995.

Trevisan, A. H., Acerbi, F., Dukovska-Popovska, I., Terzi, S., & Sassanelli, C. (2024). Skills for the twin transition in manufacturing: A systematic literature review. *Journal of Cleaner Production*, 474, 143603.

Trevisan, A. H., Arcebi, F., Dukovska-Popovska, I., Mascarenhas, J., Terzi, S., & Sassanelli, C. (2024, April). A Research Agenda Advancing Circular Economy Skills. In International Conference on Interoperability for Enterprise Systems and Applications (pp. 207-220). Cham: Springer Nature Switzerland.

Trevisan, A. H., Boscarato, A., Acerbi, F., Terzi, S., & Sassanelli, C. (2025). Enhancing Circular Economy education and training for the manufacturing sector: A holistic skills framework. *Journal of Environmental Management*, 380, 124982.

Trevisan, A. H., Lobo, A., Guzzo, D., de Vasconcelos Gomes, L. A., & Mascarenhas, J. (2023). Barriers to employing digital technologies for a circular economy: A multi-level perspective. *Journal of Environmental Management*, 332, 117437.

Van Deursen, A. J., & Van Dijk, J. A. (2014). Digital skills: Unlocking the information society. Springer.

Van Laar, E., Van Deursen, A. J., Van Dijk, J. A., & De Haan, J. (2017). The relation between 21st-century skills and digital skills: A systematic literature review. *Computers in human behavior*, 72, 577-588.

Van Laar, E., Van Deursen, A. J., Van Dijk, J. A., & De Haan, J. (2020). Determinants of 21st-century skills and 21st-century digital skills for workers: A systematic literature review. *Sage Open*, 10(1), 2158244019900176.

Vitti, M., Trevisan, A. H., Ocampo, H. R., Cuentas, V. K., Sarbazvatan, S., Terzi, S., & Sassanelli, C. (2025). A competency map for circular economy education. Procedia Computer Science, 253, 336-345.

Xu, L., Zhang, J., Ding, Y., Sun, G., Zhang, W., Philbin, S. P., & Guo, B. H. (2022). Assessing the impact of digital education and the role of the big data analytics course to enhance the skills and employability of engineering students. *Frontiers in Psychology*, 13, 974574.

Yawson, R. M., & Greiman, B. C. (2017). Strategic flexibility analysis of agrifood nanotechnology skill needs identification. *Technological Forecasting and Social Change*, 118, 184-194.

Yontar, E. (2023). Critical success factor analysis of blockchain technology in agri-food supply chain management: A circular economy perspective. *Journal of Environmental Management*, 330, 117173.

Evolving Competences through HR Projects: Knowledge Brokering Mechanisms in the SKiLL-PA Implementation Pathway

Paolo Canonico[1], Ernesto De Nito[2], Vincenza Esposito[3], Mario Pezzillo Iacono[4]
[1] University of Naples Federico II, Naples, Italy
[2] University of Salerno, Salerno, Italy
[3] University of Sannio, Benevento, Italy
[4] University of Campania "Luigi Vanvitelli", Capua, Italy

Abstract

Competence-based HRM reforms in public administrations often remain trapped in formal compliance because competence taxonomies, templates, and platform inscriptions must be translated into locally credible and administratively governable routines. This paper examines how inter-organisational knowledge brokering enables such translation in SKiLL-PA, a national, platform-mediated reform programme designed to institutionalise strategic, competency-based HRM and to connect competence modelling to workforce planning and development. We adopt a qualitative, process-oriented single-case study of UNI-1, a mid-sized public university in inland Southern Italy, drawing on interviews, documentary analysis, and direct observation. Findings show that implementation effectiveness hinged on a formally mandated academic coordinator acting as the primary knowledge broker across organisational interfaces and internal domains. Brokerage unfolded through a patterned repertoire of mechanisms: sense-giving and shared framing to reduce equivocality; translation and reframing of competence language to align behavioural competences with situated work; artefact-mediated alignment through versioned templates and traceable decision logs; cross-domain convening to integrate HR modelling with budgetary sustainability; negotiated bounded flexibility to reconcile standardisation with role hybridity; data management to secure informational integrity; legitimacy work with governance and social partners; and iterative learning routines that stabilised interpretations over time. The findings extend knowledge management research by showing how brokering work aligns interpretive, socio-technical, and institutional layers in platform-mediated competence reforms, framing inter-organisational knowledge transfer as negotiated translation rather than a top-down transmission.

Keywords: project; competences; public administration; knowledge brokering

Paper type: Academic Research Paper

1 Introduction

The expectation of new competences and responsibilities for public managers has long been a central concern in public administration scholarship and a recurring target of reform agendas aimed at strengthening administrative capacity (De Nito and Pezzillo Iacono, 2020). Yet competence-based reforms do not travel effortlessly. The knowledge infrastructures that underpin competence-centred public organisations (e.g. classification schemes, coordination mechanisms, professional standards, learning routines, and digital repositories) remain only partially integrated with HR policies and organisational design choices. The resulting misalignment sustains a persistent gap between strategic intent and operational capability, which is frequently addressed through short-term training initiatives or compliance-oriented workforce plans rather than through an organisation-wide redesign of competence architectures (Canonico et al., 2025). These implementation fragilities are well documented. Competency frameworks may be inadequately integrated with staffing, development, and reward systems; they may be confined within bureaucratic procedural compliance routines instead of serving as practical tools for managerial decision-making and learning; or they may be disputed as mechanisms of managerial control when they challenge established professional domains and evaluative norms (Horton et al., 2002). These dynamics become even more pronounced when competence frameworks are embedded in digital platforms that reconfigure what is made visible, how accountability is enacted, and how competence-related data are produced, standardised, and circulated (Nisio et al., 2023).

In complex public-sector organisational change projects initiated through top-down reform programmes, competence-based HRM reforms are not only "implemented" but *translated*: they require the circulation, reinterpretation, and stabilisation of categories and prescriptions as they travel across organisational tiers and disciplinary domains, from central programme designers and implementing partners to local administrations where the reform must be rendered workable. In practice, this translation entails aligning standardised competence artefacts with situated work arrangements, accountability expectations, and data infrastructures, through iterative negotiation among heterogeneous stakeholders and professional communities.

SKiLL-PA is a national reform initiative designed to develop an integrated approach to strategic HRM grounded in competency-based HRM, supported by a digital toolkit and an interoperability agenda across national platforms for recruitment, skills development, and integrated planning. The programme encompasses the mapping of innovative HRM practices, the co-design of a competency-based framework, and piloting and implementation in participating administrations. In this study, SKiLL-PA is treated as a sustainability-oriented competence-building project: it aims at durable competence development, institutional learning, and public value creation, while mobilising distributed expertise and coordinated action across organisational boundaries. As such, it provides a particularly informative empirical setting for examining how competence-based templates and digital artefacts are made locally workable in knowledge-intensive public organisations.

Instead of conceptualising adoption as the presence of favourable conditions, the paper examines the inter-organisational mechanisms that make the reform workable in practice, focusing on knowledge brokering activities that connect programme templates, organisational vocabularies, and operational constraints. Specifically, we focus on the role of a *knowledge broker*, an actor who connects heterogeneous communities and vocabularies, enabling knowledge transfer and translation across domains and organisations. Our objective is to explain how brokering is enacted in practice during the implementation of a competence-based, platform-mediated HR innovation: which mechanisms and practices allow the project's competence taxonomy and profiling rules to be interpreted, negotiated, adapted, and stabilised within a specific administration.

Accordingly, the paper addresses the following research question: *What mechanisms and practices do knowledge brokers mobilise to transfer and translate competence-related knowledge across organisational and disciplinary boundaries during the local customisation and implementation of SKiLL-PA?*

Empirically, we develop an in-depth case study of *UNI-1*, a mid-sized public university in inland Southern Italy, where competency-based HRM was co-designed through iterative translation into digital and procedural artefacts. Building on established conceptualisations of competence as a multidimensional construct (knowledge, skills, attitudes, and behaviours) and as a basis for competency modelling (Campion et al., 2011), we treat competence frameworks not merely as classificatory devices but as knowledge infrastructures that stabilise expectations, coordinate distributed expertise, and provide a shared language through which work design and HR decisions are negotiated across organisational boundaries.

2 Theoretical Framework: Knowledge Brokerage as Translation in Inter-organisational Projects

Collaboration between heterogeneous actors from different organisations involved in an inter-organizational change management project often involves complex knowledge transfer. All the partners have their own nomenclature, demands and expectations of innovation (Sandberg and Tsoukas, 2015). Partners have to adopt each other's perspectives and select a framework and words that can easily be understood by others.

To facilitate knowledge transfer, actors need to work together to overcome their individual and organisational interests and needs and collaborate to achieve a common goal (Canonico et al., 2025). Individuals use their own language and framework but are pushed to move from their "natural environment" (De Rond & Bouchikhi, 2004).

Specific forms of expertise play an important role in facilitating the coordination and transfer of knowledge across boundaries. According to Haas (2015) the concept of knowledge broker could be interpreted in different ways and is something different (even if with some overlap) from gatekeeper and boundary spanner. For example, knowledge brokers are intermediaries between unrelated groups or individuals geared toward knowledge gathering and dissemination. The main difference between gatekeepers and knowledge brokers is their position. Gatekeepers belong to one of the groups they link, but research presents knowledge brokers as "*liaisons*" linking two different groups without belonging to either or linking two individuals with no previous relation (Haas, 2015). On the contrary, boundary spanners could be defined as links between a unit and its environment (Cross and Prusak, 2002) who can play several different functions (Haas, 2015). In this sense when in a cross-disciplinary project academics act to put together people with different background and belonging to different organizations (or organizational units) could be typically considered as knowledge brokers more than gatekeepers or boundary spanners.

However, the concept of knowledge brokers is not unique and deserves a clarification. In a traditional perspective (Wenger, 2000) knowledge brokers are members of a community of practice who adopt different roles and practices, such as boundary spanners, roamers, outposts and pairs (Wenger, 2000) and act as brokers between communities (Ishiyama, 2016). This perspective suggests that the role of knowledge broker is critical. However, it is also a part of day-by-day work. It is therefore practiced within and outside communities and is influenced and socially constructed by the knowledge broker.

As reported by (Haas 2015 p. 1037), considering knowledge brokers as a bridge between different communities they do not belong, is very close to the role identified by Brown and Duguid's (1998) as the "translator". They are people who enable knowledge transfer across different organisational actors. They mediate between different disciplinary perspectives, facilitating communication, and creating a shared understanding of the problem or issue (Ishiyama, 2016). They act as liaison, gatekeepers, coordinators, representatives, and itinerant brokers. Recently this idea of translation has been applied in the knowledge management literature, trying to investigate the processes and tools that could help in translating knowledge (typically from a domain to another one). Donnelly (2014) has focused his attention on the role of design as a way to facilitate knowledge translation.

In our perspective, knowledge brokers should promote change, help to eliminate barriers and promote the creation of a shared language and culture. Chiambaretto et al. (2019, p. 585) stated: "*knowledge brokers are defined as actors who fulfil the role of an intermediary within the knowledge transfer process between disconnected parties (…) … They can take many forms, from consultancies to project managers*".

Perrin (2013) has investigated how knowledge brokers act to facilitate knowledge transfer inside an organization, focusing on two main aspects: the translation activities (especially related to the ability to create explicit knowledge for instance, through the generation of written documents) with concern about the capacity of filling a document creating explicit knowledge and the organizational activities (addressing the role of managers who are in charge of identifying knowledge and defining incentives to transfer knowledge and good practices inside the organization).

Our paper wants to go deeper in understanding what happens in an inter-organizational context, focusing on how knowledge brokers work to facilitate knowledge transfer. More specifically, in a cross-disciplinary project where people belong to different organizations but also to different domains. We identified the knowledge broker as someone not related to a specific community, but as someone who acts to connect people coming from different contexts, working "together" in a specific project.

3 Methodology

The study adopts a qualitative, process-oriented single-case design (Eisenhardt, 1989; Yin, 2018) aimed at explaining how knowledge is transferred and translated across organisational and disciplinary boundaries during the implementation of a competence-based, platform-mediated HR innovation. While the empirical setting is anchored in UNI-1, the analytical focus is not the organisation per se but the inter-organisational knowledge-brokering work through which SKiLL-PA's competence grammar, templates, and digital artefacts are rendered locally actionable. Accordingly, the unit of analysis is the set of brokering mechanisms and practices enacted at the interfaces among UNI-1, the programme owner (the Department), and the implementing partners (IPA and NIA), as well as across internal domains (HR/organisational development, finance, and administrative units).

In UNI-1, the academic coordinator functioned as the main knowledge broker and boundary spanner, translating the programme's competence grammar across organisational interfaces (Department/partners - University) and disciplinary domains (HRM, finance, and digital/data work).

UNI-1 was selected as an information-rich site for observing knowledge brokering in action. First, as a knowledge-intensive public organisation with structurally hybrid roles and limited slack, UNI-1 provides a context where translation work is not optional but constitutive of implementation: competence models must be adapted to multifunctional arrangements and to constrained data and governance infrastructures. Second, UNI-1's implementation architecture intentionally combined HR and finance within the core project team, increasing the likelihood of domain friction (HRM language, budgetary logics, digital/data requirements) and making the brokering function both visible and consequential.

Data collection follows a triangulation strategy combining interviews, documentary materials, and direct observation.

First, semi-structured interviews were conducted with (1) UNI-1 project-team members (HR/organisational development and finance), (2) internal governance and employee-representation actors involved in validation and sensemaking, and (3) programme counterparts where access was possible. Interviews reconstructed sequences of interactions and focused on how actors experienced translation challenges, how conflicts of interpretation were resolved, and what forms of mediation were considered legitimate.

Second, documentary analysis examined the artefacts through which knowledge was circulated, stabilised, and rendered governable. This included procedural guidance issued by the Department, templates and competence libraries provided by IPA and NIA, meeting minutes, internal communications, successive versions of UNI-1's competence-based professional system, and ancillary HR records used to validate contractual classifications and training histories.

Third, direct observation focused on the interactional settings in which brokering practices are performed: alignment meetings with programme counterparts, internal profiling and validation sessions, and workshops devoted to interpreting and refining competence descriptors. Two of the four authors acted as participant observers within the SKiLL-PA pathway, enabling access to the micro-processes of translation and mediation that are often under-specified in retrospective accounts. This included the practical work of making behavioural competences intelligible to administrative roles, reconciling HR language with budgetary constraints, and aligning toolkit affordances with established routines and data governance practices (Atkinson and Hammersley, 1995).

Confidentiality was ensured through systematic anonymisation: "SKiLL-PA" is used as a pseudonymic acronym, UNI-1 is a descriptive code, and potentially identifying institutional markers and platform labels were generalised where necessary to limit re-identification while preserving the organisational meaning of brokering mechanisms and enabling conditions.

4 SKiLL-PA as a Competence-Based HRM Reform: Aims, Design, and Toolkit Mechanisms

SKiLL-PA is embedded in Italy's National Recovery and Resilience Plan (PNRR) governance architecture and is tied to EU-facing milestones aimed at institutionalising an integrated, competence-centred approach to strategic HRM across public administrations. In this policy design, the move towards competence-based HRM is not framed as an isolated HR modernisation effort, but as a lever for strengthening administrative capacity and for enabling a more resilient public sector workforce model, whose progress must be monitored through periodic reporting obligations up to the PNRR deadline.

At its core, SKiLL-PA pursues the definition of a unitary and integrated HR strategy that re-design organisational systems towards a competency-based model shared across administrations, yet adaptable to different public-sector "comparti". This objective is operationalised through the construction of a competency-based framework, coupled with the support of an integrated data infrastructure intended to sustain comparability, aggregation, and evidence-based workforce planning. The plan logic explicitly responds to the limits of traditional job classification grounded in tasks and duties, promoting instead an HR system centred on competence typologies, professional profiles, recruitment and development pathways, and reward/career mechanisms that can be aligned with organisational redesign and strategic planning.

From an implementation standpoint, SKiLL-PA is delivered through a multi-actor arrangement led by the central public governance authority (the Department) and supported by an in-house implementation partner (*IPA*, *Implementation Partner Agency*) acting as the operational arm for project execution. In parallel, a second national-level partner (*NIA*, *National Investment Agency*) contributes to dissemination activities intended to expand uptake beyond early adopters. This division of labour is significant from an organisational perspective because it separates the design and piloting capacity required to craft a coherent competence management, from the scaling capacity required to diffuse templates, tools, and know-how across heterogeneous organisational contexts.

A distinctive feature of the programme's instrumentation is its coupling of design and diffusion with a formal monitoring regime based on KPI. The KPI reports function as a governance device through which implementation progress and reform effects are periodically assessed, enabling feedback, corrective actions, and (in the institutional framing of the PNRR) follow-up initiatives connected to integrated planning processes. The monitoring architecture is not merely evaluative: it also exerts structuring effects on participating organisations by stabilising what "counts" as progress and by shaping the reporting routines through which competence-based HRM becomes administratively actionable.

5 Implementing SKiLL-PA in UNI-1: coordination design and organisational enabling conditions

UNI-1, a mid-sized public university located in inland Southern Italy, joined SKiLL-PA after the programme's initial experimental wave involving approximately thirty public administrations. Entry into the initiative was not treated as the technical adoption of a "ready-to-use" HR toolkit. Rather, UNI-1 framed implementation as an organisational design challenge: how to render a competence-based, platform-mediated HR innovation actionable within a medium-to-small public administration characterised by limited slack resources, role hybridity, and recurrent task rotation and multifunctional teamwork.

The first implementation moves therefore concerned the architecture of coordination. UNI-1 constituted an inter-functional working group chaired by an academic with expertise in change management and HRM and composed of senior administrative representatives from the Personnel and Organisational Development area and the Financial Systems area. The inclusion of finance was consequential. It signalled that competence-based HRM was being assessed as a governance and sustainability issue, not as an HR-only reform: from the outset,

the working group sought to anticipate the cost implications of competence profiles for training investments, mobility options, career mechanisms, and internal resource allocation, thus reducing the risk that the model would later be delegitimised as fiscally unfeasible or organisationally "unfunded".

The working group operated at the intersection of intra-organisational alignment and inter-organisational coupling. Externally, it maintained structured interfaces with the Department (programme owner) and with the implementing partners (IPA and NIA), which provided procedural guidance, shared artefacts (templates, competence libraries, profiling logics), and standardisation requirements intended to preserve comparability across administrations. Internally, the group positioned itself as the translation locus where the programme's competence grammar could be rendered meaningful for the University's organisational language and work arrangements. This translation work unfolded through a staged implementation pathway, with distinct coordination mechanisms and a progressive tightening of the coupling between competence categories and certified personnel information.

During Phase 0 (approximately two months), UNI-1 prioritised onboarding and internal capability building. The core activity consisted of orientation meetings, internal training sessions supported by programme counterparts, and the stabilisation of decision rights and validation routines: who was authorised to curate personnel data, who could validate role labels, and how interpretive disagreements would be resolved. This phase also established a minimal coordination infrastructure. A restricted-access digital workspace was used to manage shared documentation, version control, and the traceability of modelling decisions—features that are particularly important where HR-related information is sensitive and where accountability demands are non-trivial. Regular virtual meetings, aligned with milestones and procedural requirements issued by the Department, created cadence and reduced equivocality, allowing the team to absorb the cognitive burden of the model without fragmenting responsibilities across organisational silos.

Phase 1 (approximately four months) marked the shift from familiarisation to modelling and implementation work. UNI-1 produced a local work plan by selecting among the methodological options offered by SKiLL-PA, thereby tailoring sequencing while remaining within the programme's core requirements. The working group then engaged in iterative cycles of data mobilisation, toolkit application, profiling, and validation as the competence framework was translated into a competence-based representation of UNI-1's professional system. In practice, this involved moving beyond the organigram as a purely structural depiction and constructing a professional system in which roles, bundles of competences, and pathways could be made comparable and usable for HR decision-making.

A pivotal operational decision concerned how to represent role plurality. Given the frequency of blended assignments and multifunctional teamwork in medium-to-small universities, the team implemented a "prevailing activity" criterion: each employee was profiled primarily against the activity that absorbed most of their time and responsibility, while secondary activities were recognised but not fully formalised in the model to avoid undermining interpretability and managerial tractability. This choice reflects a broader organisational design trade-off between descriptive completeness and governance usability. In competence-based HRM, frameworks that attempt exhaustive representation often become unwieldy and encourage compliance-driven documentation rather than operational uptake; conversely, a parsimonious representation can increase usability and the likelihood that competence profiles inform staffing, development priorities, and internal mobility discussions.

As modelling progressed, tensions emerged at the interface between standardisation and contextual fit. UNI-1 did not contest the competence-based premise; it questioned the adequacy of certain standard role labels for an organisational context in which rotation and cross-functional teamwork are routine responses to staffing constraints. Toward the end of Phase 1, internal actors requested the possibility to adapt selected elements of the general model to better represent UNI-1's division of labour. The response from programme counterparts proved enabling: rather than enforcing rigid uniformity, IPA and NIA supported a negotiated adaptation process that preserved the overarching grammar while allowing semantic adjustment of descriptors and labels. Over an additional month, an adaptation sprint, UNI-1 refined the pre-definition of its professional system within the Department's modelling framework, enhancing interpretability for local actors without compromising comparability. This episode illustrates a key enabling condition for sustainable implementation: bounded flexibility within a standardised reform, which allows local translation and reduces the risk of the model being bureaucratically contained as a short-lived compliance artefact.

The coordination regime underpinning this work combined complementary mechanisms. The restricted-access digital platform enabled distributed collaboration and auditability while protecting sensitive information flows; virtual meetings maintained inter-organisational alignment and ensured cadence; and periodic face-to-face sessions functioned as intensive translation arenas in which competence libraries were reconciled with situated knowledge of administrative work. Department-issued procedures anchored the process within

recognised institutional authority, reducing interpretive drift and providing a formal backbone for validation decisions. Within this regime, the academic chair performed a boundary-spanning and knowledge-brokering function: translating the competence grammar into locally legitimate categories, facilitating the internal legitimacy of behavioural and relational competences, and supporting communication with governance bodies and employee representatives when modelling choices required organisational endorsement.

Two enabling conditions were particularly salient in explaining why UNI-1 could complete the modelling phase within the expected timeframe and with limited resistance. The first concerns data access and informational integrity. Competence profiling required certified information on contractual profiles, role allocations, career trajectories, and training histories; involving actors with legitimate access to comprehensive personnel data reduced validation cycles and mitigated perceptions of arbitrariness. The second concerns expert boundary spanning. The academic coordinator's role strengthened coupling between programme expectations and local feasibility constraints, improved interpretive alignment when model properties were ambiguous, and enhanced legitimacy across organisational constituencies. Boundary spanning operated simultaneously as an epistemic mechanism (translation and knowledge transfer) and a political mechanism (legitimacy building and conflict mediation), which is crucial when competence-based innovations have latent implications for evaluation, mobility, and career trajectories.

By the end of Phase 1, UNI-1 produced a competence-based representation of its professional system that functioned as a new organisational artefact: it provided a shared language for discussing work, staffing, development priorities, and the alignment between HR practices and public value goals. This artefact constitutes the necessary substrate for the subsequent stage, engagement with internal governance bodies and social partners, through which the model can be institutionalised within HR routines rather than remaining confined to the project perimeter.

To consolidate the within-case analysis, Table 1 synthesises the mechanisms through which the academic coordinator enacted knowledge brokering across organisational and disciplinary boundaries during UNI-1's implementation pathway.

Table 1. Mechanisms of Brokerage in the UNI-1 SKiLL-PA Implementation Pathway

Brokerage mechanism	Core practice enacted by the broker	Boundary/interface bridged	Contribution to implementation	Where
SENSE-GIVING AND SHARED FRAMING	Established a shared interpretive frame for competence-based HRM and clarified decision rights, expectations, and milestones with internal actors and programme counterparts.	UNI-1 internal domains ↔ Department and partners	Reduced equivocality; aligned meanings and responsibilities before technical work started.	P0
TRANSLATION AND REFRAMING OF COMPETENCE LANGUAGE	Reframed behavioural and relational competences through concrete administrative episodes (interdependencies, stakeholder interactions, service continuity) to increase local intelligibility and legitimacy.	Competence grammar ↔ situated work practices	Enabled uptake of the competence library beyond formal labels; supported credible profiling.	P1
ARTEFACT-MEDIATED ALIGNMENT	Used templates, competence libraries, versioned documents, and a traceable decision log to stabilise interpretations and coordinate iterative revisions.	People ↔ artefacts across organisational boundaries	Turned discussions into auditable outputs; reduced interpretive drift and rework.	P0–P1
CONVENING CROSS-DOMAIN COORDINATION	Designed structured interactions among HR/organisation, finance, and administrative units; synchronised internal work with external procedural guidance.	HR ↔ finance ↔ operations	Integrated feasibility (budget/sustainability) with competence modelling; strengthened implementability.	P1
NEGOTIATED BOUNDED FLEXIBILITY	Brokered a negotiated adaptation of role labels/descriptors to fit hybrid work patterns while preserving the programme's core grammar and comparability requirements.	Standardisation ↔ contextual fit	Avoided bureaucratic containment; increased local usability without abandoning coherence.	P0-P1
DATA MANAGEMENT	Mobilised certified personnel data (contractual profiles, training histories, role allocations) and coordinated validation to prevent arbitrary or contested classifications.	Data governance ↔ modelling work	Shortened validation cycles; reduced resistance by grounding profiles in trusted data.	P1

Brokerage mechanism	Core practice enacted by the broker	Boundary/interface bridged	Contribution to implementation	Where
LEGITIMACY AND STAKEHOLDER MEDIATION	Translated modelling choices into governance-relevant implications (fairness, careers, development priorities) in interactions with senior leadership and employee representation.	Technical model ↔ institutional arenas	Built endorsement pathways; protected implementation from political/legitimacy shocks.	P1 end
ITERATIVE LEARNING AND FEEDBACK LOOPS	Maintained reflective routines (check-ins, issue logs, recap notes) that supported incremental refinement and learning across cycles.	Project cycles ↔ organisational learning	Sustained momentum; enabled continuous improvement of descriptors and mappings.	P0–P1

Legend: P0 = onboarding and internal capacity-building; P1 = modelling and translation work

Building on these empirically grounded mechanisms, the discussion interprets UNI-1's implementation dynamics through the paper's theoretical framework on inter-organisational knowledge transfer and brokerage.

6 Discussion and Conclusion

The SKiLL-PA case suggests that implementation hinges less on the provision of codified HR artefacts (e.g., competence libraries, profiling logics, procedural scripts, and platform features) than on an organisation's capacity to translate them into a coherent system of structures, practices, and routines, one that is forged through a complex process of negotiation and adaptation among the public administration's key stakeholders as the reform is enacted.

What crosses organisational boundaries, therefore, is not "knowledge" as an abstract object, but a bundle of categories, representations, and prescriptions that becomes consequential only once it is anchored in situated routines, accountability relations, and data infrastructures. In this sense, knowledge transfer is better understood as a socio-organisational accomplishment, produced through interpretive alignment and boundary work, rather than as a linear act of transfer (Yih-Tong Sun & Scott, 2005).

At project level, SKiLL-PA intentionally couples standardisation and customization. Tools, templates, competence taxonomies, procedural milestones, KPI-driven reporting, and a dedicated digital environment collectively provide a shared grammar meant to stabilise comparability across administrations. Yet the same grammar can intensify interpretive and political tensions when it encounters organisations whose division of labour is structurally fluid.

In UNI-1, the potential for misfit was amplified by role hybridity, task rotation, and multifunctional teamwork, features that sit uneasily with the implicit assumption, often built into competence models, that roles can be discretely bounded and consistently labelled over time. This is precisely the cross-domain condition in which actors carry distinct criteria of legitimacy, and where "making knowledge travel" depends on negotiation and translation rather than dissemination alone (Sandberg and Tsoukas, 2015). Collaboration, in turn, requires participants to step outside habitual frames and co-construct a vocabulary that can be shared and operationalised, overcoming domain-specific interests and interpretive inertia (Canonico et al., 2020).

In UNI-1, this translation burden was addressed through a formally mandated brokerage role. The academic coordinator, an internal faculty member appointed through a rectoral directive, was empowered to convene units, access required organisational information through legitimate channels and represent the University in inter-organisational alignment with the Department and programme partners. The role operated across HR, finance, and the programme interface, bridging domains whose assumptions about work, data, and legitimacy only partially overlap. This configuration aligns with accounts that conceptualise knowledge brokers as intermediaries who enable knowledge gathering and dissemination across partially disconnected groups and interpretive communities (Haas, 2015), and with the broader notion of boundary spanning as linking internal units to external environments in ways that enable coordination and knowledge flow (Cross and Prusak, 2002).

The UNI-1 implementation narrative indicates that brokering was enacted through a combination of interactional, artefactual, and institutional mechanisms, activated at critical episodes where standardisation encountered contextual constraints. First, the broker performed translation and reframing to render the competence grammar locally intelligible. In several profiling sessions, administrative staff initially treated behavioural competences as generic managerial rhetoric, peripheral to "real" administrative tasks. The coordinator re-specified these competences by anchoring them to concrete work situations (such as cross-office coordination, managing interdependencies under time pressure, and handling stakeholder expectations) thereby making them legitimate and actionable. This resembles the "translator" role emphasised in organisational knowledge accounts, whereby brokers mediate disciplinary perspectives and enable shared

understanding by reframing content into locally meaningful categories (Ishiyama, 2016). It also echoes the argument that brokering across communities of practice involves connecting practices and meanings rather than merely transmitting information (Wenger, 2000).

Second, the broker mobilised artefacts as coordination devices to stabilise meaning across interactions. The restricted-access digital workspace, versioned templates, and traceable decision logs did more than store documentation; they provided reference points that disciplined discussion, surfaced disagreements, and allowed interpretive closures to be fixed into successive artefact versions. This is consistent with the idea that boundary-crossing mechanisms rely on pragmatic infrastructures that support communication and the construction of a common collaborative framework (Rajalo and Vadi, 2017). In UNI-1, artefact mediation reduced the likelihood that the competence model would remain an episodic discourse and increased its chances of becoming a governable organisational object.

Third, brokerage became decisive when UNI-1 confronted the tension between standard labels and local work reality. A composite vignette (consistent with observed dynamics) illustrates the point. During Phase 1, the team attempted to profile a cluster of staff whose work blended procurement, digital support, and process coordination across departments. The programme template required assignment to a single standard role label that presupposed stable functional boundaries. HR staff argued that any single label would misrepresent work and generate perceived inequities; finance warned that misclassification could distort training and development priorities. The coordinator convened a "profiling clinic" with IPA/NIA counterparts and the Department's procedural referent, using three artefacts to structure negotiation: the programme role taxonomy, UNI-1 internal job descriptions and workload evidence, and a draft mapping note proposing retention of the prevailing-activity rule alongside a controlled annotation field for secondary responsibilities. The outcome was a bounded adaptation: UNI-1 preserved the programme's core profiling logic (to protect comparability) while refining descriptors and embedding a limited mechanism to represent role plurality without proliferating idiosyncratic labels. This episode illuminates a central brokerage mechanism - negotiated flexibility under a shared grammar - through which the broker converted prospective resistance into a jointly authorised adjustment, sustaining both local credibility and programme alignment. From a knowledge translation standpoint, this aligns with the view that design work can function as a translation mechanism by shaping representations so that knowledge can travel across contexts while retaining usability (Donnelly, 2014).

Fourth, the broker performed institutional and legitimacy work to protect implementation from bureaucratic containment. The formal mandate enabled early engagement with internal governance bodies and employee representatives, translating technical modelling decisions into organisationally salient stakes: recognisability of roles, perceived fairness, implications for development pathways, and the sustainability of resource commitments. This legitimacy work is consistent with definitions of knowledge brokers as intermediaries who remove barriers and foster shared language and culture in transfer processes (Chiambaretto et al., 2019). It also clarifies why UNI-1's inclusion of finance in the core team was enabling: it reduced the likelihood that the competence system would be dismissed as aspirational yet "unfunded," strengthening the credibility of the emerging professional system as a basis for managerial decisions.

Overall, the case refines the paper's argument by specifying that inter-organisational knowledge transfer in competence-based, platform-mediated reforms depends on brokerage that aligns three layers simultaneously: interpretive alignment across domains (what competences mean for local work and how they are recognised as legitimate), socio-technical alignment (how the competence grammar is inscribed into digital and documentary artefacts with attention to data integrity and traceability), and institutional alignment (how the competence system is authorised, defended, and made negotiable with internal constituencies).

The knowledge broker's dynamic role involved interdependent and overlapping functions. We identified three main functions. The first was facilitating the alignment of different languages and lenses using different brokerage mechanisms (sense-giving, translation of competence language, artefact-mediated alignment, convening cross-domain coordination, negotiated bounded flexibility, data governance, iterative learning and feedback loop, i.e. table 1). The second was connecting separated know-how, needs, and practices by managing interactions, which involved managing interactions and negotiating relationships between different stakeholders to connect knowledge, needs, and practices, and increase the permeability of boundaries between team members. The third function was sustaining engagement in HR change project by building coordination and communication mechanisms in a bottom-up approach that involved participation, involvement, and cooperation between different organisational actors.

7 Ethics Declaration

Ethical clearance was not required for this study under the authors' institutional procedures, as the research involved no interventions, no collection of special category personal data, and no foreseeable risks beyond those encountered in routine organisational activities.

8 AI Declaration

The authors used AI-based tools solely to support language editing and proofreading (e.g., improving clarity, style consistency, and correcting typographical or grammatical errors). No AI tool was used to generate, fabricate, or modify empirical data, results, or interpretations. The authors explicitly confirm that they carefully reviewed, verified, and accepted all AI-assisted edits and that they take full responsibility for the final content of the manuscript.

References

Atkinson, P. and Hammersley, M. (1995) Ethnography: Principles in Practice, 2nd ed., Routledge, London.

Brown, J. S., & Duguid, P. (1998) "Organizing knowledge", *California Management Review*, 40(3), pp. 90–111.

Campion, M.A., Fink, A.A., Ruggeberg, B.J., Carr, L., Phillips, G.M. and Odman, R.B. (2011) "Doing competencies well: Best practices in competency modeling", *Personnel Psychology*, Vol. 64, No. 1, pp. 225–262.

Canonico, P., De Nito, E., Esposito, V., Martinez, M., and Pezzillo Iacono, M. (2025) "Knowledge transfer and brokering in a public sector digital transformation project", *Knowledge Management Research and Practice*, 1–16

Chiambaretto P., Massé D., & Mirc N. (2019). "All for one and one for all?" – knowledge broker roles in managing tensions of internal coopetition: The Ubisoft case", *Research Policy*, 48(3), pp. 584–600.

Cross, R., and Prusak, L. (2002) "The people who make organizations go-or stop", *Harvard Business Review*, 80(6), 5–12.

De Nito, E., and Pezzillo Iacono, M. (2020) "Rethinking the Concept of Competencies for Public Managers". In: Decastri, M., Battini, S., Buonocore, F., Gagliarducci, F., Organizational Development in Public Administration, The Italian Way, Palgrave Macmillan.

Dingelstad, J., Borst, R. T., and Meijer, A. (2022) "Hybrid data competencies for municipal civil servants: An empirical analysis of the required competencies for data-driven decision-making", *Public Personnel Management*, 51(4), pp. 458–490.

Ehnert, I. (2009) Sustainable Human Resource Management: A Conceptual and Exploratory Analysis from a Paradox Perspective, Physica-Verlag, Heidelberg.

Eisenhardt, K.M. (1989) "Building theories from case study research", *Academy of Management Review*, Vol. 14, No. 4, pp. 532–550.

Getha-Taylor, H. (2008) "Identifying collaborative competencies", *Review of Public Personnel Administration*, pp. 28(2), 103–119. https://doi.org/10.1177/0734371X08315434

Kruyen, P. M., and Van Genugten, M. (2020). Opening up the black box of civil servants' competencies: Evidence from a survey in the Netherlands, *Public Management Review*, 22(1), pp. 118–140.

Haas, A. (2015) "Crowding at the frontier: Boundary spanners, gatekeepers and knowledge brokers", *Journal of Knowledge Management*, 19(5), pp. 1029–1047.

Hondeghem, A., Horton, S., and Scheepers, S. (2005). Modèles de gestion des compétences en Europe. Revue française d'administration publique, 4, 561–576.

Horton, S., Hondeghem, A. and Farnham, D. (2002) Competency Management in the Public Sector: European Variations on a Theme, IOS Press, Amsterdam.

Lundin, R.A., Arvidsson, N., Brady, T., Ekstedt, E., Midler, C. and Sydow, J. (2015) Managing and Working in Project Society: Institutional Challenges of Temporary Organizations, Cambridge University Press, Cambridge.

Nisio, A., Carapella, B. and De Filippis, M. (2023) "Competency Based Human Resources Management in public administrations: a multi-case study", *Azienda Pubblica*, No. 4.

Nonaka, I. (1994) "A dynamic theory of organizational knowledge creation", *Organization Science*, Vol. 5, No. 1, pp. 14–37.

Noordegraaf, M. (2000) "Professional sense-makers: Managerial competencies amidst ambiguity", *International Journal of Public Sector Management*, 13(4), pp. 319–332.

Perrin, A. (2013) "Knowledge lost in translation: The role of knowledge brokers in knowledge transfer", *International Journal of Information Technology and Management*, *12* (3/4), pp. 214–225.

Sandberg, J., and Tsoukas, H. (2015) "Making sense of the sensemaking perspective: Its constituents, limitations, and opportunities for further development", *Journal of Organizational Behavior*, 36(S1), S6–S32.

Silvius, G. (2017) "Sustainability as a new school of thought in project management", *Journal of Cleaner Production*, Vol. 166, pp. 1479–1493.

Virtanen, T. (2000) "Changing competences of public managers: Tensions in commitment" *International Journal of Public Sector Management*, 13(4), pp. 333–341.

Wenger, E. (2000) "Communities of practice and social learning systems", *Organization*, 7(2).

Yin, R.K. (2018) Case Study Research and Applications: Design and Methods, 6th ed., Sage, Thousand Oaks, CA.

Evaluating the Acceptance of GenAI in Italian Academies of Fine Arts

Francesco Filippo, Francesco Paolo Lagrasta, Barbara Scozzi
Politecnico di Bari - (DMMM), Bari, Italy

Abstract

Generative Artificial Intelligence (GenAI) is increasingly entering higher education, offering new opportunities for teaching, learning, content generation and creative experimentation. In artistic and design-oriented education, however, the acceptance of GenAI cannot be explained only through traditional efficiency-based criteria. Within Italian Academies of Fine Arts (AFAM), creative work is closely linked to identity, authorship, aesthetic judgment, expressive autonomy and cultural meaning. For this reason, understanding the factors that shape students' and educators' intention to use GenAI requires a contextualized extension of established technology acceptance models.

This paper investigates the acceptance of GenAI in Italian Academies of Fine Arts by developing and preliminarily validating a survey instrument grounded in the Technology Acceptance Model (TAM) and enriched with constructs specific to creative contexts. The study builds on a systematic literature review conducted on Scopus, which identified 274 articles at the intersection of artificial intelligence, creativity and technology adoption. A core dataset of 31 studies was then examined to identify recurrent theoretical frameworks, constructs and measurement items. Based on this analysis, TAM was selected as the reference model and extended with four context-specific constructs: trust in GenAI systems, aesthetic satisfaction, creative autonomy and subjective norm. The methodological design combines literature-based construct development, questionnaire design and qualitative validation through a cognitive interviewing approach. The questionnaire is structured around a seven-point Likert scale and includes items adapted from prior acceptance studies and reformulated for GenAI-enabled creative practices. A preliminary validation phase was conducted with participants familiar with artistic and creative domains, to assess semantic clarity, item interpretability, cognitive load and potential response biases. The paper contributes to the literature by proposing a TAM-based model tailored to creative higher education. It also offers practical implications for AFAM institutions, educators and curriculum designers interested in supporting responsible and context-sensitive GenAI integration. The findings suggest that GenAI acceptance in artistic education depends not only on usefulness and ease of use, but also on trust, perceived control over the creative process and satisfaction with the aesthetic quality of AI-generated outputs.

Keywords: Generative Artificial Intelligence; Higher Education; Technology Acceptance Model; Creative Autonomy; Academy of Fine Arts

Paper type: Academic Research Paper

1 Introduction

Generative Artificial Intelligence (GenAI) is transforming the ways in which digital content is conceived, produced and revised. Through natural language prompts, sketches or iterative interactions, users can now generate texts, images, videos, layouts and other creative outputs with unprecedented speed and accessibility. These technologies are increasingly embedded in educational environments and creative software ecosystems, where they are presented as tools for supporting ideation, experimentation, personalization and productivity (Kim, 2025; Kanont et al., 2024).

In higher education, the diffusion of GenAI has raised strategic, pedagogical and ethical questions. Institutions are required to understand how students and teachers perceive these tools, which factors influence their intention to use them, and how curricula and institutional practices should evolve in response. The relevance of this issue is reflected in the Special Track on the Acceptance of Generative Artificial Intelligence in Higher Education, which explicitly calls for studies on influencing factors of AI adoption, technology acceptance, institutional preparedness, curriculum innovation and the regulation or support of AI use in higher education. The track also highlights the relevance of models such as TAM, UTAUT, CFA and SEM-based approaches for studying these phenomena.

The acceptance of GenAI in artistic and creative higher education deserves specific attention. In Italian Academies of Fine Arts, hereafter referred to as AFAM institutions, creative production is not merely a technical activity. It is connected to identity, authorship, expressive autonomy, cultural sensitivity and aesthetic judgment. Unlike the adoption of traditional information systems, the adoption of GenAI in this context does not only depend on whether the technology is useful or easy to use. It also depends on whether students and educators

trust the system, perceive that it preserves their creative agency, and consider the generated outputs aesthetically meaningful and aligned with their intentions.

This study addresses this issue by developing a model for evaluating GenAI acceptance in AFAM institutions. The research builds on previous work on GenAI adoption in creative contexts and on a systematic literature review focused on the intersection between artificial intelligence, creativity and technology adoption. The literature review identified a corpus of 274 articles through a Scopus query and selected a core dataset of 31 papers for deeper semantic and methodological analysis. This review showed that established models such as TAM, UTAUT and TPB remain central in the study of technology adoption, but also that their application to GenAI in creative domains requires contextual adaptation.

The paper responds to a dual gap. First, technology acceptance models have often been applied to organizational, business, healthcare or general educational settings, while creative higher education remains underexplored (Cabero-Almenara et al., 2024; Zhang and Hou, 2024). Second, traditional acceptance constructs may not fully capture the symbolic, emotional and aesthetic dimensions of GenAI use in artistic practice. For example, perceived usefulness may explain whether a student believes that GenAI improves performance, but it does not fully explain whether the student feels creatively in control. Similarly, output quality may be relevant in technical contexts, but in artistic domains it is necessary to consider aesthetic satisfaction and cultural alignment.

The objective of this paper is therefore to develop and preliminarily validate a survey instrument for evaluating the acceptance of GenAI in Italian Academies of Fine Arts. The study adopts the Technology Acceptance Model as its theoretical foundation (Davis, 1989), while integrating additional constructs derived from the literature and from the specific characteristics of creative work. The model includes perceived usefulness, perceived ease of use, attitude toward use and behavioral intention, together with trust, aesthetic satisfaction, creative autonomy, facilitating conditions and subjective norm.

The main research question guiding the study is:

RQ: Which factors influence the behavioral intention to use GenAI in Italian Academies of Fine Arts, and how do creative-context constructs extend traditional technology acceptance models?

2 Relevance of the Research

The relevance of this research lies in the increasing need to understand the acceptance of GenAI in higher education and, more specifically, in creative and artistic institutions. The rapid diffusion of GenAI tools requires universities and academies to make decisions about curriculum design, student support, academic integrity, creative practice and institutional regulation. These decisions should not be based only on technological enthusiasm or fear, but on empirical knowledge about how users perceive and evaluate GenAI.

The AFAM sector represents a particularly meaningful context for this investigation. Italian Academies of Fine Arts are part of the national Higher Artistic and Musical Education system and provide first- and second-cycle degrees in fields such as painting, sculpture, scenography, design, multimedia arts and related creative disciplines. These institutions train future artists, designers, educators and creative professionals. As a consequence, their response to GenAI may influence not only current educational practices, but also future norms within creative industries.

Existing research on GenAI adoption in higher education has often focused on students' general attitudes, pedagogical beliefs, perceived usefulness, academic writing or technology readiness (Cabero-Almenara et al., 2024; Kanont et al., 2024; Mourtajji and Arts-Chiss, 2024). These contributions are important, but they rarely address the specificities of artistic education. In creative contexts, users evaluate GenAI not only as a tool that may increase efficiency, but also as an actor that can influence authorship, autonomy, originality, aesthetic coherence and emotional resonance. For this reason, a context-sensitive model is needed.

The research is also relevant from a theoretical perspective. TAM and UTAUT have provided robust frameworks for explaining technology acceptance across different settings (Davis, 1989; Venkatesh et al., 2003; Venkatesh et al., 2012). However, GenAI differs from traditional software because its outputs are probabilistic, partially unpredictable and often difficult to evaluate through purely functional criteria. In creative work, the value of an output is not simply determined by accuracy or efficiency, but by meaning, style, originality and alignment with the user's expressive goals. This makes it necessary to integrate constructs such as trust, aesthetic satisfaction and creative autonomy.

The research is also relevant for practice. Educators and institutional decision makers need tools to understand whether GenAI is perceived as a support for learning and creativity or as a threat to authorship, competence and artistic identity. A validated questionnaire can help AFAM institutions identify barriers to adoption, differences between students and educators, and areas where training, guidelines or technological

support are needed. It can also inform software designers and providers interested in developing GenAI tools that are more aligned with the needs of creative learners.

3 Contribution

This study contributes to the literature on GenAI acceptance by proposing a contextualized extension of the Technology Acceptance Model for creative higher education. Rather than replacing established acceptance frameworks, the paper builds on them and adapts them to the specific characteristics of GenAI-mediated creative practice.

3.1 Theoretical contribution

The theoretical contribution of the paper is threefold.

First, the study extends TAM to the context of GenAI in artistic education. TAM traditionally explains technology acceptance through perceived usefulness and perceived ease of use, which influence attitude toward use and behavioral intention (Davis, 1989). These constructs remain relevant, but they do not fully capture the experiential, symbolic and aesthetic dimensions of GenAI use in creative domains. By incorporating creative autonomy, aesthetic satisfaction and trust, the proposed model offers a more suitable lens for studying GenAI acceptance in AFAM institutions.

Second, the study clarifies the conceptual distinction among constructs that may otherwise overlap. For example, perceived usefulness refers to the extent to which GenAI is expected to improve creative or educational performance. Aesthetic satisfaction refers to users' evaluation of the artistic, stylistic and cultural adequacy of generated outputs. Creative autonomy refers to the perceived preservation of control, authorship and intentionality during the creative process. Trust refers to confidence in the reliability, transparency and appropriateness of GenAI systems. These distinctions respond to the need to avoid conceptual ambiguity among constructs.

Third, the paper contributes to the literature by combining literature-based item selection with qualitative validation. The questionnaire was not designed only by adapting generic TAM items, but by building a structured database of items and constructs extracted from previous studies. This database included approximately 600 items retrieved from the analyzed literature and enabled the comparison of different operationalizations of the same variables. The final items were then adapted to the language and experience of creative users.

3.2 Practical implications

From a practical perspective, the study provides a tool that can support AFAM institutions in evaluating how students and educators perceive GenAI. The questionnaire can help identify whether adoption is mainly driven by perceived usefulness, ease of use, social pressure, trust, creative autonomy or aesthetic satisfaction. This information can support targeted interventions.

For educators, the model can inform teaching strategies. If students perceive GenAI as useful but threatening to autonomy, educators may need to design activities that emphasize human control, critical selection and reflective use. If students report low trust, institutions may need to provide guidance on data sources, copyright, transparency and limitations of AI-generated outputs. If aesthetic satisfaction is low, training may focus on prompt refinement, iterative experimentation and evaluation of AI-generated alternatives.

For curriculum designers and technology providers, the model highlights the need to address GenAI not as a generic digital tool, but as a technology whose acceptance depends on usability, creative control, ethical awareness, aesthetic judgment and alignment with artistic identity.

4 Method

The study adopts a mixed and exploratory research design that combines systematic literature review, construct development, questionnaire design and qualitative validation. The methodology responds directly to the reviewers' request for a more detailed specification of the empirical strategy, sampling procedure and construct definition.

4.1 Literature review and item database

The first phase consisted of a structured Scopus review designed to capture the intersection between GenAI, creativity and technology adoption. The search was limited to English journal articles in subject areas relevant

to social sciences, computer science, business, engineering, arts, decision sciences, psychology and multidisciplinary research.

The search identified an initial corpus of 274 scientific articles. This corpus was analyzed at two levels. First, a bibliometric analysis was conducted to map temporal trends, geographical distribution, keyword co-occurrences and thematic clusters. Second, a core dataset of 31 papers was selected for deeper semantic and methodological analysis. The purpose of this second level was to identify theoretical frameworks, constructs, items, samples and methods used in the literature.

A structured item database was then created. The database collected approximately 600 measurement items from the analyzed papers, including information on the original construct, theoretical model, source and possible model extension. This database supported the identification of recurring constructs and allowed the research team to compare how similar variables were operationalized across different studies.

4.2 Selection of the reference model

The Technology Acceptance Model was selected as the reference theoretical framework because of its parsimony, clarity and suitability for studying individual-level adoption (Davis, 1989). Compared with broader models such as UTAUT and UTAUT2, TAM offers a more linear structure and allows the researcher to isolate the core drivers of acceptance: perceived usefulness, perceived ease of use, attitude toward use and behavioral intention.

However, TAM was not adopted without modification. The literature review and the specific characteristics of creative contexts suggested that the basic model should be extended. GenAI is not a deterministic tool that simply supports predefined tasks. It generates outputs that may be unpredictable, aesthetically variable and culturally sensitive. For this reason, users' acceptance may depend on additional constructs related to trust, autonomy and aesthetic evaluation.

UTAUT and UTAUT2 were also considered because they include relevant constructs such as social influence and facilitating conditions (Venkatesh et al., 2003; Venkatesh et al., 2012). In the proposed model, these dimensions are retained through subjective norm and facilitating conditions, but the overall model remains TAM-based in order to preserve parsimony and focus.

4.3 Sampling and recruitment strategy

The target population of the study consists of students and educators from Italian Academies of Fine Arts and related AFAM institutions. The empirical strategy distinguishes between two phases: a qualitative pilot validation phase and a broader quantitative survey phase.

The pilot validation phase is designed to assess the clarity, interpretability and contextual adequacy of the questionnaire. It involves participants familiar with creative and artistic practices, including students, doctoral candidates, educators or practitioners in artistic and humanistic disciplines. The purpose of this phase is not to statistically validate the factor structure of the model, but to identify semantic ambiguities, cognitive load, response difficulties and potential biases before broader administration.

The broader survey phase will target a larger sample of students and educators across multiple AFAM institutions. The recruitment strategy will follow a purposive and institutional logic. Institutions will be selected to ensure variation in geographical location, public or private status, and disciplinary orientation, including visual arts, design, multimedia arts and scenography. Within each institution, both students and educators will be invited to participate in order to compare perceptions across user profiles.

The pilot sample is therefore used only for qualitative validation and cognitive pretesting, while the broader survey will require a larger sample suitable for reliability assessment, factor analysis and, where possible, structural equation modeling.

4.4 Questionnaire design

The questionnaire is structured into two main sections. The first section collects demographic and control variables, including gender, age, educational level, role, creative discipline, previous experience with GenAI and frequency of use. These variables are important because prior experience and disciplinary background may influence perceived ease of use, trust and behavioral intention.

The second section measures the constructs included in the proposed model. Items are formulated as statements and assessed on a seven-point Likert scale, where 1 corresponds to "strongly disagree" and 7 corresponds to "strongly agree". The use of a seven-point scale is consistent with prior studies on technology acceptance and allows respondents to express nuanced levels of agreement.

The items were adapted from prior literature and reformulated for GenAI-enabled creative practice. For example, items measuring perceived usefulness refer to whether GenAI can improve the quality, speed or effectiveness of creative work. Items measuring creative autonomy refer to whether users feel that they remain in control of the creative process when using GenAI. Items measuring aesthetic satisfaction refer to whether generated outputs are perceived as aesthetically coherent and aligned with the user's expectations.

4.5 Qualitative validation through cognitive interviewing

The questionnaire was preliminarily validated through a qualitative protocol inspired by cognitive interviewing. Cognitive interviewing is used to identify sources of measurement error by investigating how respondents understand, retrieve information, judge and answer questionnaire items (Tourangeau et al., 2000; Willis, 2005; Beatty and Willis, 2007).

After completing the questionnaire, participants were invited to discuss the clarity of items, the meaning of key terms, the difficulty of answering, and the perceived adequacy of the constructs. The validation focused on four dimensions: semantic clarity, cognitive load, social desirability bias and construct completeness.

This phase allowed the research team to identify whether terms such as "creative autonomy", "aesthetic satisfaction", "trust" or "GenAI-generated output" were interpreted consistently by participants. It also helped assess whether the questionnaire language was appropriate for creative education and whether some items required reformulation.

Table 1: Research design and methodological steps

Step	Description	Output
Step 1 – Literature review	Scopus-based review on GenAI, creativity and technology adoption	Corpus of 274 articles
Step 2 – Core dataset analysis	Semantic and methodological analysis of selected studies	Core dataset of 31 papers
Step 3 – Item database construction	Extraction and classification of measurement items	Database of approximately 600 items
Step 4 – Model selection	Comparison of TAM, UTAUT and TPB	TAM selected as reference model
Step 5 – Model extension	Addition of creative-context constructs	Trust, aesthetic satisfaction, creative autonomy and subjective norm
Step 6 – Questionnaire design	Adaptation of items to GenAI and creative education	Seven-point Likert questionnaire
Step 7 – Pilot validation	Cognitive interviewing and qualitative feedback	Refined instrument for broader survey
Step 8 – Future quantitative analysis	Reliability, factor analysis, SEM or regression	Empirical validation of the model

5 Proposed Model and Preliminary Validation

5.1 Construct definitions and theoretical origin

The proposed model includes nine constructs. Four constructs derive directly from TAM: perceived usefulness, perceived ease of use, attitude toward use and behavioral intention. Two constructs are incorporated from broader acceptance frameworks: subjective norm and facilitating conditions. Three constructs are introduced or adapted for the creative GenAI context: trust, aesthetic satisfaction and creative autonomy.

Table 2: Constructs for evaluating GenAI acceptance in AFAM institutions

Construct	Definition	Theoretical origin
Perceived usefulness (PU)	The degree to which users believe that GenAI improves the effectiveness, quality or productivity of their creative work	TAM
Perceived ease of use (PEOU)	The degree to which users believe that GenAI tools are easy to learn and use	TAM
Attitude toward use (ATU)	The user's positive or negative evaluation of using GenAI in creative practice	TAM
Behavioral intention (BI)	The intention to use GenAI tools in future creative or educational activities	TAM / UTAUT

Construct	Definition	Theoretical origin
Subjective norm (SN)	The perceived social pressure from teachers, peers, institutions or creative communities to use or avoid GenAI	TAM2 / TPB / UTAUT
Facilitating conditions (FC)	The perceived availability of technical, organizational and educational support for using GenAI	UTAUT
Trust (TR)	Confidence in the reliability, transparency and appropriateness of GenAI outputs and system behavior	TAM extensions / GenAI literature
Aesthetic satisfaction (AS)	Satisfaction with the aesthetic, stylistic and cultural adequacy of AI-generated outputs	Creative-context extension
Creative autonomy (CA)	The perception of maintaining control, authorship and intentionality when using GenAI	Creative-context extension

5.2 Hypotheses and structural relationships

The conceptual model assumes that GenAI acceptance in creative education is shaped by both classical acceptance factors and creative-specific constructs. Trust is expected to influence perceived usefulness and perceived ease of use because users are more likely to consider GenAI useful and manageable when they perceive it as reliable and predictable. Aesthetic satisfaction is expected to influence perceived usefulness because creative users evaluate usefulness partly through the quality and coherence of outputs. Creative autonomy is expected to influence attitude toward use because users may develop a more favorable attitude when they feel that GenAI supports rather than replaces their creative agency.

Subjective norm is expected to influence both attitude and behavioral intention, especially in educational contexts where teachers, peers and institutional guidelines shape technology-related behavior. Facilitating conditions may influence perceived ease of use and behavioral intention by reducing technical barriers and uncertainty. Finally, perceived usefulness and perceived ease of use are expected to influence attitude toward use and behavioral intention, consistently with TAM.

Table 3: Hypothesized relationships among constructs

Hypothesis	Relationship	Rationale
H1a	Trust → Perceived usefulness	Users who trust GenAI are more likely to perceive it as useful
H1b	Trust → Perceived ease of use	Trust reduces uncertainty and supports perceived manageability
H2	Aesthetic satisfaction → Perceived usefulness	In creative contexts, usefulness depends on aesthetic adequacy
H3	Creative autonomy → Attitude toward use	Users are more favorable when GenAI preserves creative control
H4a	Subjective norm → Attitude toward use	Peer, teacher and institutional expectations shape user attitudes
H4b	Subjective norm → Behavioral intention	Social pressure may directly affect intention to use
H5a	Facilitating conditions → Perceived ease of use	Technical and educational support reduces perceived effort
H5b	Facilitating conditions → Behavioral intention	Supportive environments increase willingness to adopt
H6	Perceived ease of use → Perceived usefulness	Easier tools are more likely to be perceived as useful
H7	Perceived usefulness → Attitude toward use	Usefulness supports favorable evaluations of GenAI
H8	Perceived ease of use → Attitude toward use	Ease of use contributes to positive attitudes
H9	Attitude toward use → Behavioral intention	Favorable attitudes increase intention to use

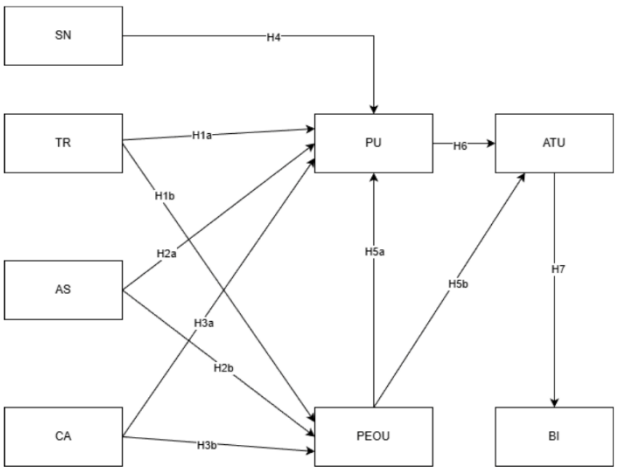


Figure 1: Proposed TAM-based model for GenAI acceptance in AFAM institutions

5.3 Questionnaire structure

The questionnaire translates the constructs into measurable items. Items were selected by comparing prior operationalizations in the item database and then adapting them to GenAI-enabled creative practice. The questionnaire includes two sections.

The first section collects demographic and contextual information. Respondents are asked to indicate their role, age, gender, educational level, creative discipline, previous experience with GenAI tools and frequency of use. These variables will be used as control variables in subsequent analysis.

The second section measures the constructs of the model through Likert-scale statements. Respondents are asked to evaluate their level of agreement with statements concerning usefulness, ease of use, trust, creative autonomy, aesthetic satisfaction, subjective norm, facilitating conditions, attitude and behavioral intention.

Table 4: Questionnaire structure and examples of item content

Section	Content	Example of measured dimension
Respondent profile	Role, age, gender, education, discipline	Student/educator profile and creative domain
GenAI experience	Prior use and frequency of use	Familiarity with text, image or video generation tools
Perceived usefulness	Impact of GenAI on creative performance	Improvement of ideation, productivity or output quality
Perceived ease of use	Effort required to learn and use GenAI	Clarity, accessibility and ease of interaction
Trust	Reliability and appropriateness of GenAI outputs	Confidence in outputs and system behavior
Aesthetic satisfaction	Alignment with aesthetic expectations	Stylistic, expressive and cultural coherence
Creative autonomy	Perceived control over the creative process	Preservation of authorship and intentionality
Subjective norm	Influence of peers, teachers and institutions	Social and educational pressure toward adoption
Facilitating conditions	Availability of support and infrastructure	Training, access to tools, institutional guidance
Attitude and intention	Overall evaluation and future use	Willingness to use GenAI in creative projects

5.4 Preliminary validation findings

The qualitative validation phase provided useful insights into the adequacy of the questionnaire. Overall, the instrument was considered relevant and aligned with the experience of creative users. Participants recognized the importance of constructs such as trust, creative autonomy and aesthetic satisfaction, confirming that GenAI acceptance in artistic contexts cannot be reduced to usefulness and ease of use.

The validation also revealed some critical issues. First, several terms required clarification. Concepts such as “GenAI”, “creative autonomy” and “aesthetic satisfaction” may be interpreted differently depending on the respondent’s background. This suggests that the questionnaire should include short preliminary definitions and examples of GenAI tools before the item section.

Second, the validation highlighted the importance of contextualizing items. Generic technology acceptance statements may be too abstract for AFAM respondents. Items should refer explicitly to creative work, artistic projects, design outputs, visual experimentation or educational activities.

Third, participants indicated that the degree of prior experience with GenAI may strongly affect responses. For this reason, previous use and frequency of use should be included as control variables. Respondents with no experience may answer based on expectations or fears, while experienced users may evaluate the technology based on actual interaction.

Fourth, the validation confirmed the risk of social desirability bias. In educational contexts, respondents may feel pressure to present themselves either as open to innovation or as protective of human creativity. For this reason, the questionnaire should emphasize anonymity and clarify that there are no right or wrong answers.

Table 5: Main insights from the qualitative validation phase

Validation area	Main issue identified	Implication for the questionnaire
Semantic clarity	Some key terms may be interpreted differently	Add short definitions and examples
Contextual relevance	Generic acceptance items may be too abstract	Reformulate items around creative and educational practices
Cognitive load	Some constructs may appear conceptually close	Clarify distinctions among usefulness, aesthetic satisfaction and autonomy
Prior experience	Users with different levels of experience may answer differently	Include previous use and frequency as control variables
Social desirability	Respondents may answer according to perceived expectations	Emphasize anonymity and neutrality
Construct completeness	TAM alone does not fully capture creative experience	Retain creative autonomy, aesthetic satisfaction and trust

5.5 Expected quantitative analysis

The refined questionnaire will be administered to a broader sample of AFAM students and educators. The quantitative analysis will first use descriptive statistics to profile respondents and summarize construct distributions. Reliability will then be assessed through Cronbach's alpha or composite reliability. Depending on the final sample size, exploratory factor analysis, confirmatory factor analysis or structural equation modeling will be used to test the measurement and structural model. If the sample is not sufficient for full SEM, hierarchical regression will be adopted to estimate the incremental explanatory power of creative-context constructs beyond traditional TAM variables. This strategy is consistent with the special track's interest in PLS-SEM, CB-SEM, CFA and technology acceptance models, while remaining appropriate for the exploratory nature of the study.

6 Discussion and Implications

The proposed model suggests that GenAI acceptance in AFAM institutions is shaped by both traditional and creative-specific factors. Perceived usefulness and perceived ease of use remain important because users are more likely to adopt technologies that support their activities and are easy to learn. However, in creative higher education these constructs are insufficient on their own.

Trust is central because GenAI systems are often perceived as opaque, probabilistic and difficult to control. Users may hesitate to rely on outputs whose sources, training data or decision logic are unclear. This is particularly relevant in artistic education, where originality and authorship are essential.

Aesthetic satisfaction is also critical. In creative contexts, users do not evaluate outputs only in terms of functionality. They assess whether outputs are stylistically coherent, expressive and aligned with their intentions. A tool may be useful in a technical sense but still rejected if its outputs are perceived as aesthetically generic or culturally misaligned. This is also consistent with responsible AI approaches that emphasize transparency, human oversight and contextual risk management (OECD, 2022; NIST, 2024).

Creative autonomy is perhaps the most distinctive construct. GenAI can be perceived as empowering when it supports exploration and expands creative possibilities. At the same time, it can be perceived as intrusive when it reduces user control or homogenizes creative expression. Therefore, acceptance depends on whether users feel that GenAI strengthens or weakens their role as authors.

7 Conclusion

This paper developed a model for evaluating the acceptance of GenAI in Italian Academies of Fine Arts. Building on TAM and related technology acceptance frameworks, the study proposed an extended model that integrates constructs specific to creative education: trust, aesthetic satisfaction and creative autonomy.

The study also responds to the need for more context-sensitive evidence on GenAI in higher education. Rather than assuming that acceptance factors are universal, it shows that local institutional settings and disciplinary cultures matter. In AFAM institutions, GenAI adoption is not only a question of performance and usability, but also of trust, control, aesthetic alignment and professional identity.

The study has limitations. The qualitative validation phase provides preliminary insights but does not yet constitute statistical validation of the full model. Future research should administer the questionnaire to a larger sample of AFAM students and educators, allowing reliability analysis, factor validation and structural model testing. Further studies may compare public and private academies, different artistic disciplines and different levels of GenAI experience.

Future developments may also support the construction of a maturity model for GenAI acceptance in creative higher education. Such a model could help institutions identify their level of readiness, design appropriate training interventions and develop policies for the responsible integration of GenAI into artistic and design education.

8 Ethics Declaration

The research design involves the administration of a questionnaire and qualitative validation activities with participants from higher education and creative contexts. Participation will be voluntary and based on informed consent. No sensitive personal data will be collected beyond information strictly necessary for the purposes of the study. Where required by the authors' institution, ethical clearance will be requested before broader data collection. Data will be anonymized and used only for academic research purposes.

9 AI Declaration

No AI tool was used to develop the paper.

References

Ajzen, I. (1991) "The theory of planned behavior", *Organizational Behavior and Human Decision Processes*, Vol. 50 No. 2, pp. 179–211.

Beatty, P.C. and Willis, G.B. (2007) "Research synthesis: The practice of cognitive interviewing", *Public Opinion Quarterly*, Vol. 71 No. 2, pp. 287–311.

Cabero-Almenara, J., Palacios-Rodríguez, A., Loaiza-Aguirre, M.I. and Andrade-Abarca, P.S. (2024) "The impact of pedagogical beliefs on the adoption of generative AI in higher education: Predictive model from UTAUT2", *Frontiers in Artificial Intelligence*, Vol. 7.

Cano, J.R. and Nunez, N.A. (2024) "Unlocking innovation: How enjoyment drives GenAI use in higher education", *Frontiers in Education*, Vol. 9.

Chan, C.K.Y. and Zhou, W. (2023) "An expectancy value theory based instrument for measuring student perceptions of generative AI", *Smart Learning Environments*, Vol. 10 No. 64.

Davis, F.D. (1989) "Perceived usefulness, perceived ease of use, and user acceptance of information technology", *MIS Quarterly*, Vol. 13 No. 3, pp. 319–340.

Gefen, D., Karahanna, E. and Straub, D.W. (2003) "Trust and TAM in online shopping: An integrated model", *MIS Quarterly*, Vol. 27 No. 1, pp. 51–90.

Goodhue, D.L. and Thompson, R.L. (1995) "Task-technology fit and individual performance", *MIS Quarterly*, Vol. 19 No. 2, pp. 213–236.

Greiner, C., Peisl, T.C., Höpfl, F. and Beese, O. (2023) "Acceptance of AI in semi-structured decision-making situations applying the four-sides model of communication: An empirical analysis focused on higher education", *Education Sciences*, Vol. 13 No. 9.

Kanont, K., Pingmuang, P., Simasathien, T., Wisnuwong, S., Wiwatsiripong, B., Poonpirome, K., Songkram, N. and Khlaisang, J. (2024) "Generative-AI, a learning assistant? Factors influencing higher-ed students' technology acceptance", *Electronic Journal of E-Learning*, Vol. 22 No. 6, pp. 18–33.

Kim, C. (2025) "Understanding factors influencing generative AI use intention: A Bayesian network-based probabilistic structural equation model approach", *Electronics*, Vol. 14 No. 3.

Mourtajji, L. and Arts-Chiss, N. (2024) "Unleashing ChatGPT: Redefining technology acceptance and digital transformation in higher education", *Administrative Sciences*, Vol. 14 No. 12.

NIST (2024) *Artificial Intelligence Risk Management Framework*. Gaithersburg, MD: National Institute of Standards and Technology.

OECD (2022) *Framework for the Classification of AI Systems*. Paris: OECD Publishing.
Schreier, M. (2012) *Qualitative Content Analysis in Practice*. London: SAGE.
Tourangeau, R., Rips, L.J. and Rasinski, K. (2000) *The Psychology of Survey Response*. Cambridge: Cambridge University Press.
van Niekerk, J., Delport, P.M.J. and Sutherland, I. (2025) "Addressing the use of generative AI in academic writing", *Computers and Education: Artificial Intelligence*, Vol. 8.
Venkatesh, V. and Davis, F.D. (2000) "A theoretical extension of the Technology Acceptance Model: Four longitudinal field studies", *Management Science*, Vol. 46 No. 2, pp. 186–204.
Venkatesh, V., Morris, M.G., Davis, G.B. and Davis, F.D. (2003) "User acceptance of information technology: Toward a unified view", *MIS Quarterly*, Vol. 27 No. 3, pp. 425–478.
Venkatesh, V., Thong, J.Y.L. and Xu, X. (2012) "Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology", *MIS Quarterly*, Vol. 36 No. 1, pp. 157–178.
Vimpari, V., Kultima, A., Hämäläinen, P. and Guckelsberger, C. (2023) "'An adapt-or-die type of situation': Perception, adoption, and use of text-to-image generation AI by game industry professionals", *Proceedings of the ACM on Human-Computer Interaction*, Vol. 7.
Wang, S.F. and Chen, C.C. (2024) "Explore the driving factors of designers' AIGC usage behavior based on SOR framework", *Frontiers in Computer Science*, Vol. 6.
Willis, G.B. (2005) *Cognitive Interviewing: A Tool for Improving Questionnaire Design*. Thousand Oaks, CA: Sage.
Zhang, W. and Hou, Z. (2024) "College teachers' behavioral intention to adopt artificial intelligence-assisted teaching systems", *IEEE Access*, Vol. 12, pp. 152812–152824.

Artificial Intelligence and Soft Skills Development in Organizations: A Systematic Literature Review

Antonio Lorena, Anna Maria Correale, Alessia Talarico, Rocco Reina
University Magna Graecia, Catanzaro, Italy

Abstract

The rapid integration of Artificial Intelligence (AI) into organizational contexts is progressively reshaping how knowledge and skills are developed and managed (Kaplan & Haenlein, 2019; Dwivedi et al., 2021). While AI-driven automation has transformed processes and decision-making, its implications for people's learning, particularly about soft skills, remain an area of growing scholarly and managerial interest. The current evidence base is still fragmented, and recent contributions have explicitly called for systematic reviews to clarify how AI supports (or constrains) soft skills (Deming, 2017; Heckman & Kautz, 2012; Balcar, 2016). Against this background, the present study adopts a Systematic Literature Review (SLR) methodology guided by the PRISMA 2020 framework (Page et al., 2021; Tranfield et al., 2003). The review focuses on peer-reviewed articles published in English, covering all available years up to February 2026. This inclusive temporal scope was adopted to provide a comprehensive view of the evolution of research on AI and soft skills development, while also incorporating the most recent contributions in the field. The search is conducted across three important academic databases: Scopus, Web of Science, and ProQuest, using the core query "Artificial Intelligence" AND "soft skills", adapted through wildcards to reflect database-specific syntax. The selection procedure follows a multi-stage screening protocol and includes a quality appraisal of the eligible studies. The review is guided by the following research question: How is Artificial Intelligence shaping the landscape of soft skills development in organizations? The synthesis is likely to highlight recurring patterns, and at the same time, expected to systematize evidence on relevant challenges such as algorithmic bias, data privacy, and the risk of dehumanizing learning contexts (Morandini et al., 2023). Overall, the study contributes by mapping the existing literature in a structured and transparent manner, while also offering actionable implications for professionals, educators, and policymakers seeking to harness AI in support of a future-ready workforce.

Keywords: Artificial Intelligence; Soft Skills; Upskilling; Reskilling; Organizational Learning

Paper type: Academic Research Paper

1 Introduction

In the contemporary organizational landscape, the rapid integration of Artificial Intelligence (AI) is transforming not only how tasks are executed, but also how knowledge and capabilities are generated, mobilised, and renewed within organizations (Kaplan et al., 2019; Dwivedi et al., 2021). AI-driven systems are increasingly embedded in decision-making, processes of communication, coordination, and learning thereby reshaping the broader architecture of work and capability development. While this transformation has been widely discussed, its implications for people's learning - particularly for the development of soft skills - remain comparatively underexplored. This is a significant gap, since soft skills such as communication, critical thinking, adaptability, and empathy are widely recognized as central to individual effectiveness, organizational resilience, and long-term employability (Heckman et al., 2012; Deming, 2017). The spread of AI appears to increase, rather than diminish, the strategic importance of these transversal capabilities. As intelligent systems become more deeply embedded in organizational routines, human actors are required not only to use new technologies, but also to interpret, contextualise, and govern their application within organizations. In this sense, soft skills are increasingly emerging as a constitutive dimension of human capital in AI-mediated workplaces. However, despite growing scholarly and managerial attention, the available evidence remains fragmented across disciplines, contexts, and methodological approaches. Recent contributions have therefore called for more systematic efforts to clarify how AI supports, reshapes, or constrains the development of soft skills across different domains of knowledge and practice (Alvarado-Bravo et al., 2024). Against this background, the present study examines the critical intersection between AI and soft skills development in organizational settings through a Systematic Literature Review (SLR) guided by the PRISMA 2020 framework (Page et al., 2021) and informed by the evidence-based review tradition in management research (Tranfield et al., 2003). By synthesizing the existing literature in a structured and transparent manner, the study aims to clarify the current state of knowledge, identify the main thematic patterns emerging from the field, and provide a stronger conceptual basis for understanding how AI is reshaping workforce development and organizational learning.

2 Methodological Design

This study adopts a SLR to consolidate and critically organize the emerging body of knowledge on the relationship between Artificial Intelligence (AI) and soft skills development in organizational settings. A systematic literature review (SLR) is highly appropriate as the research field is still fragmented and rapidly evolving, because it enables a transparent, replicable, and cumulative synthesis of the evidence (Tranfield et al., 2003; Lorena, 2018; Xiao and Watson, 2019). In line with this logic, the review was designed according to the PRISMA 2020 framework, which provides a robust protocol for study identification, screening, eligibility assessment, and final inclusion (Page et al., 2021). The review addresses the following research question: How is Artificial Intelligence shaping the landscape of soft skills development in organizations?

2.1 Research STRATEGY

The literature search was conducted in Web of Science, Scopus, and ProQuest, selected because of their breadth, multidisciplinary coverage, and relevance for management and organizational research. The search was based on the core query *"artificial intelligence"* AND *"soft skills"*, applied consistently across the three databases with minor syntax adaptations where required by each platform. To ensure comparability and methodological consistency, the search was restricted to peer-reviewed journal articles published in English, covering all available years up covering all available years up to February 2026. Detailed search settings for each database are reported below.

Table 1 - Search Strategy and Retrieval Parameters Across Databases

Database	Search options
Web of Science	Search: topic Document type: article Language: English Date range: all years to February 2026 Citation indexes: available Web of Science export set
Scopus	Search: article title, abstract, keywords Document type: article Language: English Date range: all years to February 2026
ProQuest	Search: title, abstract, and subject terms Document type: article Source type: scholarly journals Language: English Date range: all years to February 2026

2.2 Inclusion and exclusion criteria

Explicit inclusion and exclusion criteria were defined in advance to ensure relevance and analytical consistency. Studies were included published in peer-reviewed scientific journals, written in English, that addressed the relationship between AI and soft skills as their main topic, and that provided empirical, conceptual, or review-based insights relevant to organizational, work, and educational contexts. We excluded studies that did not examine the interaction between AI and soft skills, that were not part of organizational contexts, and that were not deemed sufficiently relevant after evaluating the title, abstract, or full text. Contributions developed in educational settings were retained only when they offered mechanisms, technologies, or developmental insights transferable to organizational learning and workforce development.

2.3 Data collection, extraction and synthesis

The study selection process followed the logic of the PRISMA 2020 workflow, moving from identification to screening, eligibility assessment, and final inclusion (Page et al., 2021). After retrieval from the three databases, records were checked for duplication and screened through titles and abstracts. Potentially relevant studies were then assessed in full text against the predefined eligibility criteria, leading to a final corpus of 33 articles.

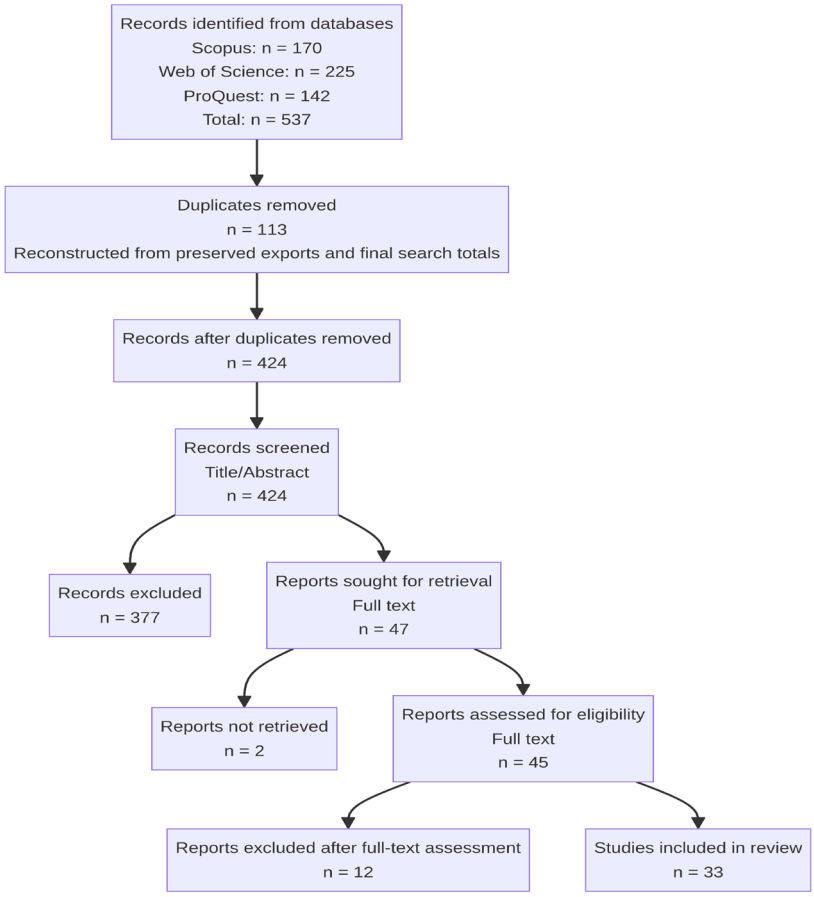


Figure 1 - PRISMA 2020 Flow Diagram of the Study Selection Process

Once the final set of studies had been identified, a structured data extraction process was conducted for each article. The extracted information included research focus, AI technologies or systems, soft skills addressed, methodological approach, principal findings, and organizational implications. The evidence was then synthesized through two complementary stages. First, a descriptive analysis was conducted to map the temporal distribution and internal composition of the corpus. Second, a thematic synthesis was developed to identify the main recurring dimensions across the reviewed studies. In line with the research objective, the synthesis was organized around three macro-areas:

- AI tools and systems used to support soft skills development;
- reported soft skills and learning outcomes;
- organizational strategies and enabling conditions, and barriers associated with AI-supported workforce development.

3 Descriptive Analysis

This section provides a concise descriptive overview of the 33 studies included in the review. Its purpose is to map the empirical structure of the corpus and highlight the main descriptive regularities, without anticipating the more interpretive discussion developed in the following section.

3.1 Publications per year

The temporal distribution of publications shows that research on AI and soft skills development in organizational and workforce-related settings is still relatively recent but has expanded markedly in the last two years of the observation period. While the earliest contributions in the final corpus appear only sporadically, the literature becomes much more concentrated from 2024 onward, with a clear acceleration in 2025. This pattern suggests that the topic has moved from a peripheral concern to a more recognizable research area, likely reflecting both the broader diffusion of AI across work contexts and the growing interest in its implications for human capability development.

Figure 2 – Annual Distribution of the Reviewed Publications (N = 33)

3.2- Main thematic synthesis aligned with the review objective

To organize the corpus descriptively, the reviewed studies were grouped into three broad thematic dimensions aligned with the review objective. The first concerns AI tools and systems used to support soft skills development and includes contributions focusing on machine learning, generative AI, chatbots, virtual and extended reality, natural language processing, and profiling systems used for training, coaching, learning support, or assessment. The second concerns learning outcomes related to soft skills. Among the most frequently mentioned in the corpus are communication, leadership, problem solving, critical thinking, empathy, adaptability, and collaboration. The third concerns organizational strategies, enabling conditions, and barriers, including reskilling projects, training design, privacy, bias, transparency, and broader implementation challenges.

Table 2 – Thematic Synthesis of the Reviewed Literature

Synthesis dimension	**Core analytical focus**	**Main recurrent elements in the dataset**	**Representative studies**
AI tools and systems used to support soft skills development	Identification and categorization of the AI technologies and system types described in the corpus.	Recurring technologies include machine learning and predictive analytics, generative AI and large language models, chatbots, natural language and speech-based systems, virtual and extended reality solutions, robotics and automation, and profiling or psychometric systems. Across the dataset, these tools are used mainly for training, coaching, learning support, decision support, assessment, evaluation, testing, and personalization.	Bartolome et al. (2025); Budnarowski et al. (2025); Ciaschi et al. (2024); Ebert et al. (2024); Tusquellas et al. (2025)
Reported soft skills and learning outcomes	Analysis of the competencies most frequently targeted and of the outcomes associated with AI-supported interventions.	The corpus recurrently refers to communication, leadership, problem solving, critical thinking, creativity, empathy, adaptability, collaboration, teamwork, negotiation, active listening, emotional intelligence, and time or stress management. Reported outcomes include personalized feedback, repeated practice, immersive or scenario-based learning, self-reflection, perceived preparedness, knowledge retention, and behavioral improvement in specific training environments.	AlAfnan et al. (2024); Bhargava et al. (2021); Budnarowski et al. (2025); Gonzalez-Rico et al. (2024); Luo et al. (2021)

Synthesis dimension	Core analytical focus	Main recurrent elements in the dataset	Representative studies
Organizational strategies, enabling conditions, and barriers	Synthesis of the organizational conditions that support or constrain the integration of AI into workforce development initiatives.	The reviewed studies recurrently associate implementation with reskilling and upskilling programs, on-the-job mentoring and coaching, human oversight, context-specific training design, and structured feedback loops. The synthesis also foregrounds challenges emphasized in the review framing, including algorithmic bias, data privacy, transparency, measurement validity, and the risk of de-humanizing learning contexts when AI-mediated development is not balanced with human interaction. Underexplored areas include longitudinal evidence, governance arrangements, and the organizational conditions under which AI-supported soft skills development can be both scalable and context-sensitive.	Bhargava et al. (2021); Bobitan et al. (2024); Bukartaite et al. (2023); Gafni et al. (2024); Sposato (2025)

Overall, this descriptive structure shows that the literature is not organized solely around the technological promise of AI, but also around the conditions under which its use in soft skills development becomes credible, scalable, and relevant for organizations.

3.3 Descriptive patterns across themes

Two broad descriptive patterns emerge from the corpus. First, the literature is strongly oriented toward the technological and developmental uses of AI, with a particular concentration on simulation, coaching and assessment. Second, the studies do not focus only on benefits, but also frequently emphasize organizational conditions and implementation challenges. As a result, the corpus is structured not merely around what AI can do, but also around the conditions under which its use in soft-skills development becomes credible and relevant for organizations.

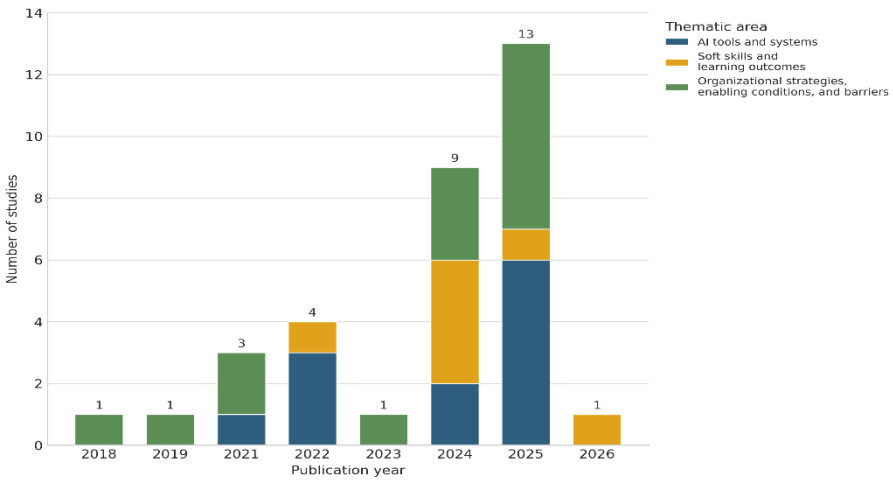


Figure 3 – Annual Distribution of Studies by Thematic Area (N = 33)

4 Results, Findings, and Discussion

The 33 studies reviewed here suggest that the literature on AI and soft-skills development is moving beyond a generic concern with technological adoption and toward a more analytically relevant question: how AI reshapes the organizational conditions through which human capabilities are developed, enacted, assessed, and governed. The descriptive analysis showed that the corpus clusters around three connected dimensions: the AI tools and systems used to support soft-skills development, the soft skills and learning outcomes most frequently addressed, and the organizational conditions that enable or constrain implementation. Building on that foundation, this section identifies three recurring and interrelated findings. First, AI is increasingly seen as a development infrastructure, not simply a technology. Second, the spread of AI appears to intensify, rather than

diminish, the strategic importance of purely human soft skills. Third, the development value of AI-based systems remains tied to organization, trust, and human-centred governance. Taken together, these patterns indicate that AI is not simply being added to existing training systems; it is contributing to a broader reconfiguration of how organizations identify, cultivate, evaluate, and legitimate human capability in hybrid work settings.

4.1 AI as a developmental infrastructure

A first major finding is that a substantial part of the literature no longer treats AI simply as a background driver of skill change or as a generic digital aid to training. Many studies frame AI as a sociotechnical infrastructure that organizes the conditions under which soft skills development occurs. From this perspective, AI plays a developmental role by diagnosing skill gaps, personalizing learning paths, providing iterative feedback, and monitoring progress over time. This infrastructural role is especially explicit in studies on AI-enabled coaching, profiling, assessment, and adaptive learning systems, where AI is placed within the architecture of development itself rather. (Luo et al., 2021; Gafni et al., 2024; Tusquellas et al., 2025).

4.1.1 Recurrent developmental functions

This first pattern becomes clearer when one examines the recurring developmental functions attributed to AI across the corpus. First, personalized developmental pathing emerges as a key function, through which AI identifies gaps, profiles learners or employees, and recommends differentiated developmental routes rather than standardized interventions. Second, coaching and feedback represent another recurring function, with AI-supported systems providing immediate and context-sensitive guidance during or immediately after performance episodes (Budnarowski et al., 2025; Luo et al., 2021; Balachandran, 2025). Third, immersive simulation and scenario-based practice appear prominently, especially in studies showing how AI-enhanced virtual or mixed environments can be used to rehearse communication, empathy, conflict resolution, or cross-cultural interaction in safe and repeatable settings (Budnarowski et al., 2025; Carvalho et al., 2025; Budnarowski et al., 2025). Finally, continuous assessment and diagnostic profiling emerge as an important infrastructural layer, particularly in studies where behavioural, linguistic, or interactional signals are translated into developmental information that can inform tailored interventions (Celsi et al., 2022; Ciaschi & Barone, 2024; Gafni et al., 2024). Across these studies, AI is valuable because it makes development more programmable, adaptive, scalable, and data-informed.

4.1.2 Conceptual significance and boundary conditions

This pattern is stronger than the claim that AI merely supports training. In the strongest studies, AI helps determine who needs what, under which conditions, and through which sequence of practice, feedback, and monitoring developmental progress can be sustained over time. This interpretation resonates with broader scholarship on organizational learning and knowledge processes, which has long emphasized that learning is processual and distributed by the infrastructures through which knowledge is created, shared, and applied (Nonaka, 1994; Tsoukas, 2009; Jarrahi et al., 2023). At the same time, some studies offer direct evidence from work contexts, while others illustrate transferable mechanisms without fully demonstrating long-term organizational effects (Bartolomé et al., 2025; Jung & Suh, 2024; Ottaviano, 2026). The literature therefore supports the claim that AI is increasingly framed as developmental infrastructure, but it also suggests that the conceptual maturity of this logic still runs somewhat ahead of its longitudinal validation.

4.2 The strategic revaluation of human soft skills

A second important finding emerging from the reviewed literature is that the spread of AI increases the strategic importance of capabilities that remain relational, interpretive, context-sensitive, and normatively oriented. This pattern is visible in conceptual, across foresight-oriented and organizational studies, which converge on the view that communication, empathy, collaboration, creativity and critical thinking continuous reskilling become even more important as AI and automation reshape work systems (Ghislieri et al., 2018; Bhargava et al., 2021; Bobitan et al., 2024). Related insights appear in studies of future skills, which consistently treat soft skills not as residual traits left untouched by technology but as capabilities that become more valuable precisely because AI is entering organizational processes (Bukartaite & Hooper, 2023; Depoo et al., 2025; Woods & Jones, 2025).

4.2.1 Core capabilities emphasized in the corpus

The reviewed studies recurrently refer to communication, leadership, problem solving, critical thinking, creativity, empathy, adaptability, collaboration, teamwork, negotiation, emotional intelligence, active listening, and time or stress management. These capabilities appear not only in papers focused on changing skill demand or workforce transition, but also in studies examining AI-enabled developmental interventions more directly, including work on mentoring, coaching, communication support, simulation-based learning, and managerial development (Luo et al., 2021; AlAfnan et al., 2024; González-Rico & Lluch Sintes, 2024). In this sense, the literature points not simply to the persistence of soft skills, but to their strategic revaluation.

4.2.2 Human–AI complementarity

This pattern aligns with a growing body of work on human–AI complementarity. Rather than framing organizational intelligence as a zero-sum contest between humans and machines, recent scholarship has emphasized the value of hybrid intelligence and human–AI symbiosis, in which analytical speed is combined with human judgment, contextual understanding, and social sensemaking (Jarrahi et al., 2023; Dellermann et al., 2019). The review findings are consistent with this perspective. Communication is important because AI outputs must be translated and made meaningful within organizational interaction. Critical thinking is important because AI-generated suggestions require evaluation rather than passive acceptance. Adaptability, empathy and collaboration are important because hybrid work systems continue to depend on affective understanding and coordinated action. The reviewed corpus therefore suggests that soft skills are becoming less peripheral and more constitutive of organizational capability in AI-mediated settings.

4.3 Organizational mediation, trust, and human-centred governance

A third finding, and arguably the one that prevents the review from sliding into technological optimism, is that the developmental potential of AI remains heavily conditioned by organizational mediation, trust, and governance. Across the corpus, the strongest interventions are rarely those in which AI replaces human developmental roles outright. Instead, they are typically hybrid arrangements in which AI complements coaching, mentoring, leadership judgment, or human interpretation. This pattern is especially clear in studies showing that AI-enabled development is most effective when embedded in broader systems of guidance and support rather than treated as an autonomous solution (Luo et al., 2021; Sposato & Dittmar, 2025; Woods & Jones, 2025). The same logic is reinforced by organizational studies emphasizing that human oversight, contextual sensitivity, and emotional attunement remain essential when AI enters capability development and workforce transformation (Ghislieri et al., 2018; Bhargava et al., 2021; Bhargava et a., 2021).

4.3.1 Governance pressures

The importance of mediation is closely tied to the governance burdens created by AI-enabled development systems. The same features that make AI attractive for people development, also make these systems governance intensive. Across the corpus, recurring concerns include privacy, bias, fairness, transparency, monitoring, trust, and the risk of dehumanizing (Bhargava et al., 2021; Ciaschi & Barone, 2024; Gafni et al., 2024). Also, olther studies on immersive and simulation-based environments, which point to additional barriers related to cost, usability, implementation, and ethical acceptability (Carvalho et al., 2025; Tutu et al., 2025; Balachandran, 2025). The broader literature on trust in AI show that successful integration depends not only on technical performance, but also on reliability, transparency, and the perceived legitimacy of the socio-technical system in which AI is embedded (Gafni et al., 2024).

4.3.2 Measurement and interpretive risk

A further issue concerns measurement. Several studies rely on behavioural proxies or inferred representations of soft skills, that enhance visibility, comparability, and scalability, yet they also risk reducing relational, judgment-based, and context-sensitive capabilities to what can be captured computationally. This tension is visible in studies on stealth assessment and NLP-based inference, where the analytical power of AI is evident but the possibility of over-instrumentalizing soft skills remains unresolved (Celsi et al., 2022; Bartolomé et al., 2025; Tutu et al., 2025). For this reason, the review does not support an understanding of AI as an autonomous answer to soft-skills development. Rather, it suggests that AI if combining structured practice, developmental analytics, human interpretation, and ethically robust governance.

4.3.3 Integrative interpretation

These findings suggest that the literature is converging around a coherent, though still maturing, understanding of AI and soft-skills development. First, AI is increasingly treated as developmental infrastructure rather than as a mere background technology. Second, this infrastructural transformation appears to revalue human soft skills by making more central to organizational performance and adaptation. Third, the developmental value of AI remains contingent on organizational mediation, trust, and governance. Read together, these findings say: AI shapes soft-skills development in organizations through the infrastructuring of development, the strategic revaluation of human capabilities, and the governance of hybrid developmental systems. All of this constitutes the main contribution of the review.

5 Conclusions and Future Research

This review set out to examine how AI is shaping soft-skills development in organizations. The evidence synthesized across the 33 reviewed studies suggests that AI should not be understood merely as a support technology for training, nor as a force that simply displaces human capability. Rather, the literature points to a more complex transformation. AI is increasingly functioning as a developmental infrastructure that reorganizes learning conditions through personalization, feedback, simulation, and analytics; at the same time, its diffusion appears to increase the strategic relevance of distinctively human capabilities such as communication, judgment, empathy, collaboration, and adaptability. Yet these potentials do not unfold automatically. Their developmental value depends on organizational mediation, hybrid human–AI arrangements, and governance mechanisms capable of preserving trust, legitimacy, and contextual sensitivity.

5.1 Main contribution

The review makes three contributions. First, it systematizes a fragmented literature by identifying a coherent set of developmental functions through which AI operates in relation to soft skills. Second, it shows that the rise of AI has been accompanied by a revaluation of human capabilities, rather than by their straightforward marginalization. Third, it clarifies that the effective integration of AI into soft-skills development is a socio-technical and governance challenge. In this sense, the review contributes to current debates on organizational learning, reskilling, and HR transformation by offering an integrative perspective that connects technological infrastructures, human capabilities, and working conditions.

5.2 Future research directions

The review points to important directions for future research. More longitudinal studies are needed to assess whether AI-enabled developmental gains persist over time rather than remaining short-term or context-bound. Additional research should examine sectoral variation, since the organizational conditions under which AI supports soft-skills development may differ substantially across industries, professions, and institutional environments. Finally, Future scholarship must shift from merely asking if AI supports soft-skills development to exploring when and how it does so effectively. Ultimately, research should focus on the specific organizational conditions that ensure AI-driven growth remains both impactful and meaningfully human.

Ethics declaration

This study is based exclusively on a systematic literature review of published sources and did not involve human participants, personal data, or confidential organizational data. Therefore, no formal ethical clearance was required.

AI declaration

AI tools were used only for language enhancement and editorial support.

References

Alvarado-Bravo, N., Aldana-Trejo, F., Duran-Herrera, V., Rasilla-Rovegno, J., Suarez-Bazalar, R., Torres-Quiroz, A., Paredes-Soria, A., Gonzales-Saldaña, S.H., Tomás-Quispe, G. and Olivares-Zegarra, S. (2024) "Artificial intelligence as a tool for the development of soft skills: A bibliometric review in the context of higher education", International Journal of Learning, Teaching and Educational Research, Vol. 23, No. 10.

Babashahi, L., Rosa, C.E., Lima, Y., Lyra, H.S., Argôlo, M., de Almeida, M.A. and de Souza, J.M. (2024) "AI in the workplace: A systematic review of skill transformation in the industry", Administrative Sciences, Vol. 14, No. 6, Article 127.

Balachandran, S. (2025) “AI-Powered Soft Skills Training: Measuring the Efficacy of VR Simulations for Cross-Cultural Competence in Global Tech Teams”, International Journal of Scientific Research & Engineering Trends, Vol. 11, No. 4, July-August, [online].

Balcar, J. (2016) "Is it better to invest in hard or soft skills?", The Economic and Labour Relations Review, Vol. 27, No. 4, pp. 453–470.

Bartolomé, J., del Río, I., Martínez, A., Aranguren, A., Laña, I. and Alloza, S. (2025) “Game On: Exploring the Potential for Soft Skill Development Through Video Games”, Information, Vol. 16, No. 10, p. 918.

Bhargava, A., Bester, M. and Bolton, L. (2021) "Employees' perceptions of the implementation of robotics, artificial intelligence, and automation (RAIA) on job satisfaction, job security, and employability", Journal of Technology in Behavioral Science, Vol. 6, pp. 106–113.

Bobitan, N., Dumitrescu, D., Popa, A.F., Sahlian, D.N. and Turlea, I.C. (2024) "Shaping tomorrow: Anticipating skills requirements based on the integration of artificial intelligence in business organizations - A foresight analysis using the scenario method", Electronics, Vol. 13, No. 11, Article 2198.

Brynjolfsson, E. and McAfee, A. (2014) The second machine age: Work, progress, and prosperity in a time of brilliant technologies, W.W. Norton & Company, New York.

Budnarowski, D., Jereczek, D., Detka, K. and Wieczorek, I. (2025) "Application of artificial intelligence and virtual reality in soft skills training with modeled personality", Applied Sciences, Vol. 15, No. 16, Article 9067.

Bukartaite, R. and Hooper, D. (2023) "Automation, artificial intelligence and future skills needs: An Irish perspective", Journal of Management Development, Vol. 42, No. 4, pp. 312–326.

Caputo, F., Cillo, V., Candelo, E. and Liu, Y. (2019) "Innovating through digital revolution: The role of soft skills and big data in increasing firm performance", Management Decision, Vol. 57, No. 1, pp. 203–223.

Carvalho, S.A., Conceição, E.S. and Marques, I.C.P. (2025) "The impact of virtual reality on employee training and learning in organisations: A systematic literature review", Applied Sciences, Vol. 15, No. 1, Article 363.

Ciaschi, M. and Barone, M. (2024) "Exploring the role of artificial intelligence in assessing soft skills", in Proceedings of the 19th Conference on Computer Science and Intelligence Systems (FedCSIS), pp. 573–578, IEEE.

Davenport, T.H. and Prusak, L. (1998) Working knowledge: How organizations manage what they know, Harvard Business School Press, Boston.

Deming, D.J. (2017) "The growing importance of social skills in the labor market", The Quarterly Journal of Economics, Vol. 132, No. 4, pp. 1593–1640.

Depoo, L., Hajerová-Mullerová, L., Kronberger, Z., Říhová, G., Stříteský, M., Hořáková, M., Legnerová, K., Palíšková, M., Němec, O., Šmíd, D., Jurčík, T. and Kopecký, M. (2025) "The impact of AI implementation on job transformation and competency requirements: Prioritising reskilling and soft skills development", Quality Innovation Prosperity, Vol. 29, No. 2, pp. 71–92.

Dwivedi, Y.K., Hughes, L., Ismagilova, E. et al. (2021) "Artificial intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy", International Journal of Information Management, Vol. 57, Article 101994.

Ebert, C. and Hemel, U. (2024) "Grow your artificial intelligence competence", IEEE Computer, Vol. 57, No. 9, pp. 144–150.

Emma, O. (2026) "Extended reality in the digital age: A literature review on technology-driven learning environments", Computers & Education: X Reality, Vol. 8, Article 100127.

Frey, C.B. and Osborne, M.A. (2017) "The future of employment: How susceptible are jobs to computerisation?", Technological Forecasting and Social Change, Vol. 114, pp. 254–280.

Gafni, R., Aviv, I., Kantsepolsky, B., Sherman, S., Rika, H., Itzkovich, Y. and Barger, A. (2024) "Objectivity by design: The impact of AI-driven approach on employees' soft skills evaluation", Information and Software Technology, Vol. 170, Article 107430.

Ghislieri, C., Molino, M. and Cortese, C.G. (2018) "Work and organizational psychology looks at the fourth industrial revolution: How to support workers and organizations?", Frontiers in Psychology, Vol. 9, Article 2365.

González-Rico, P. and Lluch Sintes, M. (2024) "Empowering soft skills through artificial intelligence and personalised mentoring", Education Sciences, Vol. 14, No. 7, Article 699.

Grama, B. and Todericiu, R. (2025) "The evolution of skill dynamics in the context of the future of work", Studies in Business and Economics, Vol. 20, No. 2, pp. 142–154.

Heckman, J.J. and Kautz, T. (2012) "Hard evidence on soft skills", Labour Economics, Vol. 19, No. 4, pp. 451–464.

Jarrahi, M.H., Askay, D., Eshraghi, A. and Smith, P. (2023) "Artificial intelligence and knowledge management: A partnership between human and AI", Business Horizons, Vol. 66, No. 1, pp. 87–99.

Jung, D. and Suh, S. (2024) "Enhancing soft skills through generative AI in sustainable fashion textile design education", Sustainability, Vol. 16, No. 16, Article 6973.

Kaplan, A.M. and Haenlein, M. (2019) "Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence", Business Horizons, Vol. 62, No. 1, pp. 15–25.

Kassa, B.Y. and Worku, E.K. (2025) "The impact of artificial intelligence on organizational performance: The mediating role of employee productivity", Journal of Open Innovation: Technology, Market, and Complexity, Vol. 11, No. 1, Article 100474.

Kitchenham, B. and Charters, S. (2007) Guidelines for performing systematic literature reviews in software engineering, Technical Report EBSE-2007-01, Keele University, Keele.

Lomsadze-Kutchava, M., Tsereteli, T. and Kutateladze, K. (2025) "Soft skills in management using artificial intelligence", Georgian Scientists, Vol. 7, No. 2, pp. 37–44.

Lorena, A. (2018) "The relation between corporate social responsibility and bank reputation: A review and roadmap", European Journal of Economics and Business Studies, Vol. 4, No. 2, pp. 7–21.

Luo, X., Qin, M.S., Fang, Z. and Qu, Z. (2021) "Artificial intelligence coaches for sales agents: Caveats and solutions", Journal of Marketing, Vol. 85, No. 2, pp. 1–18.

Morandini, S., Fraboni, F., De Angelis, M., Puzzo, G., Giusino, D. and Pietrantoni, L. (2023) "The impact of artificial intelligence on workers' skills: Upskilling and reskilling in organisations", Informing Science, Vol. 26, pp. 39–68.

Nonaka, and Takeuchi, H. (1995) The knowledge-creating company: How Japanese companies create the dynamics of innovation, Oxford University Press, New York.

Nonaka, I. (1994) “A Dynamic Theory of Organizational Knowledge Creation”, Organization Science, Vol. 5, No. 1, pp 14–37.

Ottaviano, E. (2026) “Extended Reality in the Digital Age: A Literature Review on Technology-Driven Learning Environments”, Computers & Education: X Reality, Vol. 8, p. 100127.

Osei Tutu, D., Habibiabad, S., Van den Noortgate, W., Saldien, J. and Bombeke, K. (2025) "When action speaks louder than words: Exploring non-verbal and paraverbal features in dyadic collaborative VR", Sensors, Vol. 25, No. 17, Article 5498.

Page, M.J., McKenzie, J.E., Bossuyt, P.M. et al. (2021) "The PRISMA 2020 statement: An updated guideline for reporting systematic reviews", BMJ, Vol. 372, n71.

Pinos Ullauri, L.A., Lebis, A., Karami, A., Vermeulen, M., Fleury, A. and Van den Noortgate, W. (2025) "A recommender system of postgraduate courses based on soft skills: A psychometric-inspired approach", International Journal of Artificial Intelligence in Education, Vol. 35, pp. 2117–2153.

Celsi, L., Cevallos Moreno, J.F., Kieffer, F. and Paduano, V. (2022) "HR-specific NLP for the homogeneous classification of declared and inferred skills", Applied Artificial Intelligence, Vol. 36, No. 1, Article 2145639.

Santana, M. and Díaz-Fernández, M. (2023) "Competencies for the artificial intelligence age: Visualisation of the state of the art and future perspectives", Review of Managerial Science, Vol. 17, pp. 1971–2004.

Shumeiko, N.V. and Osadcha, K.P. (2024) "Application of artificial intelligence in higher education institutions for developing soft skills", Journal of Physics: Conference Series, Vol. 2871, No. 1, Article 012027.

Sposato, M. and Dittmar, E.C. (2025) "The AI-powered future of digital transformation: Enhancing organizations and leadership development", Journal of Work-Applied Management, advance online publication.

Tranfield, D., Denyer, D. and Smart, P. (2003) "Towards a methodology for developing evidence-informed management knowledge by means of systematic review", British Journal of Management, Vol. 14, No. 3, pp. 207–222.

Tsoukas, H. (2009) “A Dialogical Approach to the Creation of New Knowledge in Organizations”, Organization Science, Vol. 20, No. 6, pp 941–957.

Tusquellas, N., Santiago, R. and Palau, R. (2025) "Professional development analytics: A smart model for Industry 5.0", Applied Sciences, Vol. 15, No. 4, Article 2057.

Woods, A. and Jones, B. (2025) "Soft skills in the age of AI: Conceptualizing the AI+EQ framework", Journal of Ethics in Entrepreneurship and Technology, advance online publication.

Xiao, Y. and Watson, M. (2019) "Guidance on conducting a systematic literature review", Journal of Planning Education and Research, Vol. 39, No. 1, pp. 93–112.

A Twin Transformation Maturity Model for Public Administrations: a Design Science Research Study

Michela Colacicco[1], Pierpaolo Pontrandolfo[1], Luigi Ranieri[2],Barbara Scozzi[1]
[1] Politecnico di Bari, Bari, Italy
[2] Università LUM Giuseppe Degennaro, Casamassima, Italy

Abstract

Digital transformation and sustainability transformation originate as different domains and have often been treated separately in both the public and private sectors. In recent years, their potential synergies, referred to as twin transformation, have begun to emerge. In this new field of research, initial contributions have primarily focused on the private sector, where a maturity model has also been developed. However, despite these early contributions, research on twin transformation remains limited, with little consideration of the public sector, and particularly local public administration. To date, the scientific literature has not identified any maturity model that specifically measures the maturity of both transformations simultaneously in public administration. To address this gap, the study adopts a design science research protocol to adapt a maturity model originally developed for the private sector to the context of local public administration. The research process includes a literature review and the design and development of the model. The results consist of a capability maturity model encompassing six capability dimensions, four maturity levels, and 99 dynamic capabilities. The model is complemented by a self-assessment instrument that enables policymakers to independently evaluate their level of twin transformation maturity. Although preliminary, these findings contribute to a more integrated understanding of digital and sustainability transformations within public administrations. By doing so, the study extends prior research on twin transformation by identifying the dynamic capabilities required to progress along the twin transformation journey.

Keywords: twin transformation, maturity assessment, public administration, dynamic capabilities.

Paper type: Academic Research Paper

1 Introduction

In recent years, the synergistic potential between digital transformation and sustainability transformation has gained increasing attention (European Commission 2022; Christmann et al., 2024), opening up a new field of research that warrants further investigation (Ologeanu-Taddei et al., 2025). To harness such potential, Christmann et al. (2024) align the digital and sustainability perspectives by defining twin transformation as "a value-adding interplay between digital and sustainability transformation efforts that improve an organization by leveraging digital technologies for enabling sustainability and leveraging sustainability for guiding digital progress" (p.495). The study further demonstrates the importance of dynamic capabilities for the implementation of twin transformation, thereby identifying 16 specific twin dynamic capabilities. Building on this foundation, Breiter et al. (2024) developed and validated the Twin Transformation Capability Maturity Model (TTCMM), a maturity model tailored to the private sector. Despite this contribution, limited attention has been devoted to the public sector, and more specifically to local public administrations. To date, the extant body of literature in this sector has proposed only a few maturity models, and these have treated the two transformations separately (Iannacci et al., 2025; Liu et al., 2022). Against this backdrop, the objective of this study is to develop a twin transformation maturity model tailored to the context of local public administrations.

To address the research objective, the study adapts a model originally designed for the business context (Breiter et al., 2024) to the public administration sector and designs and develops a self-assessment instrument that allows public administrations to measure their digital and sustainability transformation efforts in an integrated way. To achieve this, the study employs a design science research protocol following the guidelines proposed in Peffers et al. (2007). Such methodology is suitable for the development and adaptation of artifacts, especially in the case of a maturity model (Pöppelbuß and Röglinger, 2011).

2 Background

This section discusses the applicability of dynamic capabilities to the public administration sector and presents the reference model for the study.

2.1 Dynamic capabilities in public administration

Dynamic capabilities are the organizational and strategic routines through which firms achieve new configurations of resources in the market (Eisenhardt and Martin, 2000). These capabilities drive the creation, evolution, and recombination of resources into new sources of competitive advantage (Henderson and Cockburn, 1994; Teece et al., 1997). The concept of dynamic capabilities was originally developed for the private sector; however, the academic debate on the role of these capabilities has also been extended to the public sector (Panagiotopoulos et al., 2023). Public sector organizations adopt processes similar to those of private firms. These processes include the dissemination of best practices and the use of organizational resources and routines aimed at implementing policy initiatives and delivering services (Bryson et al., 2007; Pablo et al., 2007). Examples of dynamic capabilities in public administration include the ability to learn new practices, information gathering and evaluation, information-sharing systems, intellectual capital, and performance management routines (Vignieri and Zeinali, 2023).

Although in the private sector Christmann et al. (2024) have only recently identified sixteen twin transformation dynamic capabilities, there is still limited evidence in the public sector, and more specifically in the public administration. At present, dynamic capabilities in public administration are addressed in the literature mainly in relation to digital transformation and sustainability transformation as separate domains (Vignieri and Zeinali, 2023; Bhatia et al., 2025).

2.2 Reference model

Breiter et al. (2024) developed, evaluated, and demonstrated the Twin Transformation Capability Maturity Model. The model encompasses six dimensions, namely strategy and leadership, culture and employees, ecosystem and partnerships, products and services, operations, and technology, and four maturity levels: (1) awareness of combining digitalization and sustainability, (2) twin transformation development, (3) twin transformation implementation, and (4) true twin transformer. Developed for the business sector, the model is considered the starting point for the development of a twin transformation maturity model specifically tailored to the public administration context. It is regarded suitable for achieving the research objective. In particular, among its strengths, it assigns equal importance to both transformations and acknowledges that organizations may possess varying levels of capabilities related to both digital transformation and sustainability transformation.

3 Methodology

The design science research methodology was adopted to address the research questions, following the seven steps proposed by Peffers et al. (2007), as illustrated below.

Step 1: Identification and motivation of the problem. The first step was to identify and justify the problem. To this end, the literature was analyzed in order to identify the scientific contributions related to the development and application of maturity models in the digital and/or sustainability domains in public organizations. No model was identified that specifically measured the maturity of both transformations simultaneously in the public administration. Furthermore, given the limitations of existing maturity models in public administrations (e.g., published in non-indexed journals and lack of validation in real-world contexts), it was deemed appropriate to extend the analysis to the private sector. To this end, the only model identified (Breiter et al., 2024) that measured the maturity of twin transformation was selected. Based on the results of the analysis, the lack of twin maturity models specifically tailored for public administration was identified as a problem. Such a problem was confirmed by discussions with staff and the director of a local public administration in Italy.

Step 2: Defining the objectives of a solution. Defining the objectives of the solution represented a crucial step in guiding the adaptation and development of the artifact (Peffers et al., 2007). To address the problem, an existing maturity model developed for the private context (Breiter et al., 2024) was adapted to align with the specific characteristics of public administrations. Specifically, three solution objectives were outlined.

Solution objective 1. To adapt the maturity model originally conceived for the private sector to the specificities of public administration.

Solution objective 2. To redefine the dimensions and capabilities of the model from a public value perspective.

Solution objective 3. To develop the maturity measurement instrument, in line with the recommendation of Breiter et al. (2024), in order to provide a self-assessment instrument accompanying the model.

In light of the differences between the private sector and public administration, primarily in terms of objectives and the capabilities intended to be developed, we identified the first two solution objectives. Their achievement required an analysis of the characteristics of the public administration context, its purposes, the regulatory framework, and governance logic, in order to ensure relevance and full applicability. The third solution objective was linked to a limitation of the original model (de Bruin et al., 2005). Although it was demonstrated in a real context, the maturity assessment was guided by the authors. The model lacked a self-assessment tool that would enable organizations to independently carry out a maturity evaluation.

Step 3: Design and development. After identifying the problem and defining the objectives, the artifact was designed and developed. We began the design of the maturity model by mapping the capability dimensions and subsequently the dynamic capabilities, both in relation to the dimensions and the maturity levels. The details of the design are reported below.

Regarding capability dimensions, culture and employees, operations and technology remained unchanged. Human resources, technological resources and culture were all essential elements for public administration, particularly within the domain of digital transformation (Vignieri and Zeinali, 2023; Aras and Büyüközkan, 2023). The culture and employees dimension was understood as the study of the effects of the change generated by the digital transformation process on the existing culture (Aras and Büyüközkan, 2023) and the presence of appropriate skills and capabilities within the organization (Vignieri and Zeinali, 2023). For the operations dimension, the value creation processes through design techniques remained valid in the context of public administration. In particular, value creation was understood as a collaborative and multi-actor process through which different stakeholders participate in the design, delivery, and evaluation of services, generating benefits that are not limited to the individual experience but also include social, economic, and institutional outcomes (Osborne et al., 2022). The technology dimension was defined as the availability of technological infrastructures enabling people to adopt methods and tools to collect, analyze, and use data in support of decision-making processes (Vignieri and Zeinali, 2023). The other three capability dimensions of the original model underwent some adjustments. The strategy and leadership dimension was expanded to include the concept of governance. Governance was identified as a fundamental aspect in the public sector in order to ensure the successful implementation of digital transformation and its sustainability (Aras and Büyüközkan, 2023). The ecosystem and partnership dimension was renamed external ecosystem and refers to the management of external interactions and how such interactions were transformed into value. The product and service dimension was renamed public services, which play a central role for public administration, also in relation to sustainability and digital transformation (Gudmundsdottir et al., 2024; Ly, 2025).

Digital and sustainability dynamic capabilities were identified in the literature with reference to the public sector. The literature resulted in 164 dynamic capabilities, which were then compared with those proposed in the model by Breiter et al. (2024). Points of overlap were identified, such as the need to equip employees with new capabilities to face digital and sustainability transformations, and the importance of establishing partnerships to facilitate the exchange of information and the achievement transformation objectives. The overlapping capabilities remained almost unchanged in meaning and were slightly improved in form. Other capabilities were adapted more extensively. For instance, the original capability influencing legislators on twin transformation standards was expanded to the following: contributing to the definition of standards and guidelines for twin transformation through collaboration among the government, central agencies, regional authorities, autonomous provinces, and local administrations. Another example is the original capability focus on cradle-to-cradle approaches, which was adapted to focus on the entire life cycle of services and their economic, environmental, and social sustainability.

The discussion on the capabilities to be adapted, added, and removed was conducted among the authors until consensus was reached. All adaptations in terms of dimensions and capabilities with respect to the original model were systematically considered and documented. The comparison and adaptation process was also complemented by meetings with officials of an Italian public administration.

A self-assessment instrument was designed and developed to accompany the model. This instrument allowed public organizations to independently assess their maturity and was based on a four-point scale: Fully Implemented (FI), Largely Implemented (LI), Partially Implemented (PI), and Non Implemented (NI). Using scales for these purposes improved the reliability and consistency of responses and map outcomes to different levels of maturity (de Bruin et al., 2005). To design the instrument, each capability of the model was transformed into an interrogative statement. Respondents were asked to rate each statement on a scale from FI to NI. The maturity level was measured as a function of the number of capabilities that have been implemented. To reach a level, all statements must be evaluated as FI or LI. A level was not considered achieved if even a single statement within that level was assessed as PI or NI (DOE, 2022). For example, to attain level 3, all statements

related to levels 1, 2, and 3 had to be evaluated as FI or LI. The maturity level achieved was displayed through seven graphs.

Step 4&5: Evaluation and demonstration. The evaluation phase is ongoing. At present, ten experts have been selected and have been contacted by e-mail through the authors' contacts. Some of the experts were selected for their knowledge of the public administration sector, others for their relevance in the field of academic research, especially related to the development of maturity models. A first meeting was conducted, and two additional meetings were planned. To carry out the evaluation, following the meeting participants were asked to complete a questionnaire designed to gather quantitative and qualitative data. The questionnaire is structured into two sections. The first section included questions based on a 5-point Likert scale, for which "1" represents "strongly disagree" and "5" represents "strongly agree," to assess the model and instrument in terms of fidelity with real world phenomena, completeness, level of detail, internal consistency, ease of use, and usefulness (Davis, 1989; March and Smith, 1995). The second section included open-ended questions aimed at collecting feedback and suggestions for improving the maturity model. Based on the evaluation, the model will be refined as necessary. Subsequently, the model will be applied in a real-world setting.

Step 6: Communication. The design process of a maturity model will be documented and communicated in a way understandable for academics, public managers and practitioners (Becker et al., 2009; de Bruin et al., 2005).

4 Discussion

In this section, the preliminary results of the study are presented and discussed, and their implications are outlined.

4.1 Twin transformation capability maturity model

The model is grounded in the need, identified in the literature and through engagement with public administration staff, to assess the concurrent efforts undertaken by public administrations in both the digital and sustainability domains. Existing maturity models do not address these two domains simultaneously and are often developed for the private sector. This study extends prior research by adapting a private-sector model to the public administration sector, thereby providing an understanding of the capabilities required to progress toward a twin transformation journey. At the current stage of the study, the result is a theoretical model, developed through a rigorous design process, which provides insights into the evolution of dynamic capabilities for twin transformation. Although preliminary, these results suggest a possible integration of the two transformations also for public administrations, which are thus beginning to manage the twin transformation by leveraging the mutual benefits of digitalization and sustainability, in line with Christmann et al. (2024).

The proposed maturity model comprises six dimensions, namely strategy, leadership and governance, culture and employees, external ecosystems, public services, operations and technology. Each dimension encompasses a specific set of dynamic capabilities, which are organized across the four maturity levels, as listed in Table 1. At the first level, organizations are fully developed in one of the two domains and begin to develop awareness of the other domain. This initial level is characterized by capabilities related to a single transformation (digital or sustainability), as well as capabilities that enable the initial experimentation of integrated digitalization and sustainability initiatives. The second level of maturity assumes that organizations have embarked on a path toward the integration of the two domains; accordingly, they start to design capabilities that enable a joint consideration of both domains. For example, they develop a transformation strategy that incorporates both sustainability and digitalization, consider their opportunities and threats, and invest in twin projects that jointly enable the achievement of digital and sustainability objectives. At the third level, organizations operationalize the twin transformation strategy for organization-wide implementation. This is achieved, for example, through the implementation of new processes and services and the provision of training to employees. At the fourth level, organizations incorporate the principles of twin transformation into decision-making processes and are committed to continuously optimizing and aligning the two transformation objectives, also beyond organizational boundaries, by establishing dialogue with other local authorities.

Table 1: Twin Transformation Maturity Model: dimensions, capabilities and levels.

Capability dimension	1 Awareness of combining digitalization and sustainability (DT)	1 Awareness of combining digitalization and sustainability (ST)	2 Twin transformation development	3 Twin transformation implementation	4 True twin transformer
Strategy, leadership and governance	Understand the value and challenges associated with the sustainability of digital transformation; Translate the values of digitalization into concrete actions; Be aware of technological developments and trends assessing associated opportunities and threats; Structure a governance model for digital transformation that includes guidelines for cybersecurity; Allocate human and financial resources to digitalization initiatives; Promote the use of digital technologies	Understand the value and challenges related to the digitalization of sustainability transformation; Translate the values of sustainability into concrete actions; Be knowledgeable about approaches and methods related to sustainability; Structure a governance model for sustainability transformation that includes guidelines for cybersecurity; Allocate human and financial resources to sustainability initiatives; Promote life cycle analysis of services and their associated economic, environmental, social, and territorial sustainability; Adopt tools for the planning, monitoring, and evaluation of public policies from a sustainability perspective	Develop a strategy that fosters sustainability, ethical, and inclusive development through innovation and digitalization in service of individuals, communities, and territories; Strengthen investments that jointly enable the achievement of digital and sustainability objectives; Establish a performance measurement system with indicators to monitor digitalization and sustainability in an integrated manner; Foster the advancement of the twin transformation through top management commitment; Develop long-term infrastructure innovation plans that take sustainability impact into account; Evaluate and identify opportunities and threats associated with the twin transformation; Perceive and recognize the needs of internal and external users to guide decision-making from a twin transformation perspective	Promote the implementation of transformation initiatives through (top) management; Reconfigure the value proposition and organizational identity to enhance efforts toward the twin transformation	Evaluate the achievement of twin transformation objectives and, where necessary, revise transformation strategies accordingly; Promote pathways of continuous process improvement; Evaluate the organization's digital and sustainability performance to inform the planning of future interventions
Culture and employees	Train human capital in relation to digital competencies; Encourage employees and promote a culture of innovation; Develop competencies related to the use of digital technologies within one's area of activity	Train human capital in relation to sustainability competencies; Encourage employees and promote a culture of innovation; Develop competencies related to the application of sustainability principles within one's area of activity	Foster creativity and self-fulfillment to stimulate digital and sustainability innovation; Strengthen and introduce principles and values capable of fostering cultural change aligned with a digital and sustainability vision; Facilitate the exchange of ideas regarding transformation through open discussions among managers; Create an environment that stimulates employees; Enhance leadership and supervisory competencies of managers	Train employees on administrative, digital, ecological, and social transformation topics; Redefine employee and (top) management incentive systems in relation to digital and sustainability performance; Enable new ways of working and new tools introduced as a result of the twin transformation; Improve coordination within the organization; Promote employee engagement on issues related to the twin transformation	Evaluate employee satisfaction with the training provided and their competencies in terms of sustainability and digitalization; Evaluate employee and (top) management incentive systems in relation to digital and sustainability performance

Capability dimension	1 Awareness of combining digitalization and sustainability (DT)	1 Awareness of combining digitalization and sustainability (ST)	2 Twin transformation development	3 Twin transformation implementation	4 True twin transformer
External ecosystem	Increase transparency in administrative decision-making through e-democracy tools; Interact with different stakeholders on the themes and implementation of digital transformation in the public administration; Organize moments of discussion with other public administrations aimed at monitoring compliance with digital transformation requirements	Foster dialogue with partners to share sustainability goals and practices; Interact with different stakeholders on the themes and implementation of sustainability transformation in the public administration; Organize moments of discussion with other public administrations aimed at monitoring compliance with sustainability transformation requirements	Co-design and co-plan with stakeholders to achieve joint projects; Define processes for the exchange of information and knowledge with stakeholders; Define mechanisms to identify priorities among the interventions emerging from stakeholder engagement	Integrate partners' knowledge and competencies on the twin transformation, developed through established dialogue; Implement processes for the exchange of information and knowledge with stakeholders; Initiate dialogue with stakeholders on potential issues arising from the implementation of twin transformation pathways	Contribute to the definition of standards and guidelines for the twin transformation through collaboration among the Government, central authorities, Regions, Autonomous Provinces, and local authorities
Public services	Promote the development of digital services; Promote projects aimed at making territories smart and digital (e.g., smart cities)	Promote the collection of environmental and social data through the creation of a (potentially digital) platform; Promote projects aimed at making territories sustainability	Redesign or develop services in accordance with the objectives of the twin transformation (e.g., sustainability procurement policies through digitalization); Reorganize service development teams by including specialists in sustainability and digitalization; Establish systems that facilitate knowledge management to develop and improve public services; Consider needs emerging from interactions with users to foster innovation in the development of new services and projects; Identify the economic and financial, social, and environmental benefits and sustainability of the services implemented	Implement services in accordance with the objectives of the twin transformation; Implement new processes and solutions, including through the use of emerging technologies and artificial intelligence, in an ethically conscious and sustainability manner; Implement systems that facilitate knowledge management to develop and improve public services	Assess economic, environmental, and social sustainability throughout the entire life cycle of services and projects implemented

Capability dimension	1 Awareness of combining digitalization and sustainability (DT)	1 Awareness of combining digitalization and sustainability (ST)	2 Twin transformation development	3 Twin transformation implementation	4 True twin transformer
Operations	Promote the use of integrated digital systems for information exchange and operational management; Design and implement transformation through the reengineering of processes using digital technologies; Establish routines to minimize waste in operational processes and improve sustainability performance (e.g., energy efficiency of public buildings)	Promote the introduction of criteria for digitally assessing the sustainability performance of operations; Design and implement transformation through the reengineering of processes, potentially also using digital technologies; Promote interdepartmental collaboration by encouraging the sharing of operational practices and information to achieve common results; Recognize the need to protect, maintain, and conserve resources supporting public infrastructures (e.g., managing resources through a green approach); Orient the selection towards the procurement of goods and services with environmental characteristics; Recognize the need to maintain and improve health quality	Design a real-time data collection and monitoring system (for sustainability data); Design new processes and advanced technologies oriented toward the twin transformation; Implement processes to continuously and effectively share information across the organization; Strengthen the capacity to acquire external funding in a twin transformation perspective; Plan, evaluate, and acquire new equipment to support operations from a twin transformation perspective	Implement new processes and advanced technologies oriented toward the twin transformation; Implement processes, IT resources, and services to improve their effectiveness in terms of sustainability	Evaluate advanced data analytics tools, including AI-based solutions, to continuously improve activities and operational performance related to the twin transformation; Monitor and improve systems that support the updating, sharing, and reorganization of knowledge within the organization
Technology	Establish data management mechanisms to support decision-making processes; Scouting and acquisition of digital technologies for the reengineering of processes; Promote awareness of sustainability within the IT unit (green IT / efficient IT); Pursue the modernization of infrastructures for the delivery of digital services that respond to users' needs	Establish data management mechanisms for sustainability/ESG-related data collection to ensure their usability in decision-making processes; Scouting and acquisition of digital technologies for the reengineering of processes from a sustainability perspective	Design technological solutions, including potentially AI-based systems, that are human-centric, reliable, secure, and sustainability; Design energy-efficient and economically sustainability digital infrastructures	Facilitate inter-organizational (inter-office) cooperation through the adoption of shared digital tools; Promote the use of advanced analytics in management dashboards to monitor the achievement of twin transformation objectives; Ensure the effective use of data and information analysis tools to enhance organizational knowledge assets and steer the organization toward a data-driven model	Continue to leverage the sustainability potential of emerging digital technologies, consolidating already implemented projects

4.2 Implications for research

The study has significant implications at the theoretical, practical, and societal levels. The theoretical contribution of the study is threefold. First, it contributes to the existing body of literature on twin

transformation, a field that has been predominantly explored within the private sector, by developing a twin transformation maturity model for local public administrations, thereby addressing a significant research gap. Second, the self-assessment instrument helps operationalize the maturity evaluation, bridging the gap between theory and practice. Third, the study informs theory by examining the application of dynamic capabilities to twin transformation in public administration.

At the practical level, the proposed model provides public managers with a self-assessment instrument based on multiple dimensions, capabilities, and maturity levels. Such an instrument enables the assessment of the current level of twin transformation maturity, highlighting the capabilities and, consequently, the strategic decisions, resources, and investments that need to be designed in order to advance toward twin transformation.

At the societal level, the study fosters dialogue and awareness of the concept of twin transformation in the public sector, thereby influencing public debate and promoting a more informed discussion of its implications.

5 Conclusion

The study adapts and refines an existing model to the context of local public administration by incorporating new dynamic capabilities and expanding its dimensions. The resulting maturity model is structured around six dimensions and 99 dynamic capabilities and helps organizations simultaneously measure their efforts in the digital and sustainability domains. The model is also accompanied by a self-assessment instrument that enables policymakers to independently evaluate their maturity, representing a novel feature compared to previous models. However, the study is not without limitations. At its current stage, the model is conceptual and is based on a framework originally developed for the private sector. Although significant adaptations were made, some underlying assumptions may not be fully transferable to public administrations. In the future, expert interviews and real-world case applications are planned to verify the model's full applicability and its alignment with the public administration sector.

6 Ethics declaration

Participants were informed about the purpose of the study and their contribution. The data, which did not contain sensitive information, was treated confidentially and anonymity was guaranteed.

7 AI declaration

No artificial intelligence tools were used in developing the paper.

References

Aras, A., and Büyüközkan, G. (2023) "Digital transformation journey guidance: a holistic digital maturity model based on a systematic literature review", Systems, 11(5), 213.

Becker, J., Knackstedt, R. and Pöppelbuß, J. (2011) "What makes a useful maturity model? A framework of general design principles for maturity models and its demonstration in business process management", European Conference on Information Systems (ECIS), Helsinki, Finland, Vol. 19, June.

Bhatia, R., Pihlajamaa, M., Hyytinen, K., and Jäntti, A. (2025) "Capabilities for transformative sustainability management in cities", Public Management Review, 27(4), 1179–1203.

Breiter, K., Crome, C., Oberländer, A.M. and Schnaak, F. (2024) "Dynamic capabilities for the twin transformation climb: A capability maturity model", Information Systems Frontiers, Vol. 26, pp. 2205–2226.

Bryson, J.M., Ackermann, F., and Eden, C. (2007) "Putting the resource-based view of strategy and distinctive competencies to work in public organizations", Public Administration Review, 67(4), 702–717.

Christmann, A.S., Crome, C., Graf-Drasch, V., Oberländer, A.M. and Schmidt, L. (2024) "The TT butterfly: Capabilities for an integrated digital and sustainability transformation", Business & Information Systems Engineering, Vol. 66, No. 4, pp. 489–505.

Davis, F.D. (1989) "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology", MIS Quarterly, Vol. 13, No. 3, September, pp. 319–340.

de Bruin, T., Freeze, R.D., Kulkarni, U., and Rosemann, M. (2005) "Understanding the main phases of developing a maturity assessment model", Proceedings of the Australasian Conference on Information Systems (ACIS).

Department of Energy United States (2022) Cybersecurity Capability Maturity Model.

Eisenhardt, K.M. and Martin, J.A. (2000) "Dynamic capabilities: What are they?", Strategic Management Journal, Vol. 21, No. 10/11, pp. 1105–1121.

European Commission (2022) 2022 Strategic Foresight Report, European Commission, Brussels.

Gudmundsdottir, S., Sigurjonsson, T.O., Hlynsdottir, E.M., Fridriksdottir, S.D., and Ingibjargardottir, I.S. (2024) "Sustainable digital change: the case of a municipality", Sustainability, 16, 1319.

Henderson, R., and Cockburn, I. (1994) “Measuring competence? Exploring firm effects in pharmaceutical research”, Strategic Management Journal, 15, 63–84.

Iannacci, F., Karanasios, S., Viscusi, G., McManus, R., Rupietta, C. and Tan, C.W. (2025) “Unboxing maturity models: A set-theoretic perspective on e-Government configurations over time”, Journal of Strategic Information Systems, Vol. 34, Article 101874.

Liu, J., Wang, C., Chen, N., Chen, Z., Xu, L., Du, W. and Zhang, Y. (2022) “Towards sustainable smart cities: Maturity assessment and development pattern recognition in China”, Journal of Cleaner Production, Vol. 370, Article 133248.

Ly, B. (2025) “Leveraging leadership and digital transformation for sustainable development: insights from Cambodia’s public sector”, Sustainable Futures, 9, Article 100545.

March, S.T. and Smith, G.F. (1995) “Design and natural science research on information technology”, Decision Support Systems, Vol. 15, pp. 251–266.

Ologeanu-Taddei, R., Hönigsberg, S., Weritz, P., Wache, H., Mittermeier, F., Tana, S., Pekkola, S. (2025) “The relationship of DT and corporate sustainability: Synergies and tensions”, Technological Forecasting and Social Change, Vol. 210, Article 123809.

Osborne, S.P., Powell, M., Cui, T. and Strokosch, K. (2022), “Value Creation in the Public Service Ecosystem: An Integrative Framework”, Public Admin Rev, 82: 634-645.

Pablo, A.L., Reay, T., Dewald, J.R. and Casebeer, A.L. (2007) “Identifying, enabling and managing dynamic capabilities in the public sector”, Journal of Management Studies, Vol. 44, No. 5, pp. 687–708.

Panagiotopoulos, P., Protogerou, A., and Caloghirou, Y. (2023) “Dynamic capabilities and ICT utilization in public organizations: An empirical testing in local government”, Long Range Planning, 56(1), Article 102251.

Peffers, K., Tuunanen, T., Rothenberger, M.A. and Chatterjee, S. (2007) “A design science research methodology for information systems research”, Journal of Management Information Systems, Vol. 24, No. 3, pp. 45–77.

Spaltini, M., Acerbi, F., De Carolis, A., Terzi, S., and Taisch, M. (2024) “Toward a technology roadmapping methodology to enhance sustainable and digital transition in manufacturing”, Production & Manufacturing Research, Vol. 12, No. 1, Article 2298572.

Tabares, S., Parida, V., and Chirumalla, K. (2025) “Twin transition in industrial organizations: Conceptualization, implementation framework, and research agenda”, Technological Forecasting and Social Change, Vol. 213, Article 123995.

Teece, D.J., Pisano, G., and Shuen, A. (1997) “Dynamic capabilities and strategic management”, Strategic Management Journal, 18(7), 509–533.

Vignieri, V., and Zeinali, Z. (2023) “Dynamic capabilities as paths to organizational readiness for digital transformation: an analysis of the NRRP horizontal reform ‘public administration digitalization’”, Azienda Pubblica, 36(3), 271–310.

Experiences of Knowledge-based Human Resources Management in the Finnish Public Sector

Sini Tenhovuori, Nina Helander
Tampere University, Tampere, Finland

Abstract

Knowledge-based human resource management (KHRM) functions as an enabler of sustainability by making visible the consequences of HR practices and by supporting evidence-informed decision-making that extends beyond financial indicators to incorporate human, social and environmental impacts. Accordingly, sustainable and strategic KHRM creates the conditions for balancing service effectiveness, employee well-being and long-term societal responsibility.
The purpose of this study is to explore how KHRM is implemented in Finnish public sector organisations through the experiences of employees involved in HRM. Aim is to identify the organisational, managerial, cultural and competence-related factors that either enable or hinder the implementation and development of strategic and sustainable KHRM. The empirical data was collected from 14 Finnish public sector organisations.
According to the findings of this study, the development and implementation of KHRM in Finnish public organisations are specifically promoted by a set of clearly identifiable enabling factors. The need to strengthen KHRM has been widely recognised, creating also a shared mandate for development. The integration of KHRM into organisational strategies, together with explicit support from top management, provides a strong structural and leadership foundation for its advancement. Furthermore, an organisational culture that is receptive to data-driven practices, the increasing interest in HR indicators, and a general orientation toward knowledge-based development all act as an important driver that actively facilitate and reinforce the implementation of KHRM. However, the study found many factors that hinder the development of KHRM. These included, for example, challenges in the availability of the HR data, quality of HR information as well as ambiguities related to HR concepts and indicators. In addition, the implementation of KHRM was weakened by deficiencies in the utilization of human resources, deficiencies in the processes of KHRM, manual reporting, fragmented data sources.
KHRM is slowly developing in Finnish public sector organisations. Identifying factors that are enabling or hindering the implementation and the development of KHRM helps organisations to strengthen the maturity of KHRM.

Keywords: Knowledge management, Human resources management, HR-analytics, public sector, management

Paper Type: Academic Research Paper

1 Introduction

Knowledge-based management has been recognised as a foundation for both strategic and sustainable human resources management, as it enables organisations to utilise data, information and knowledge in a more systematic and forward-looking manner (Partanen et al. 2023; Vakkuri et al. 2023). In the Finnish public sector, the provision services is fundamentally dependent on human resources. The availability of skilled, sufficient and well-being personnel is therefore a prerequisite for delivering high-quality and effective public services (Keva 2025; Laihonen & Kokko 2023). At the same time, personnel costs constitute a substantial share of the public sector expenditure (Kauhanen 2025).

From a KM perspective, the public sector represents particularly relevant research context due to its strong requirements for transparency, accountability and effectiveness, as well as its exposure to complex and often competing societal expectations (Partanen et al. 2023; Erkkilä 2020). Prior research has shown that KM in the public sector operates within a context characterised by institutional complexity, where multiple logics, actors and governance mechanisms shape how data and information are produced, interpreted and utilised (Laihonen & Kokko 2023). While earlier studies have emphasised the strategic importance of KM for public organisations, evidence from recent research suggests that knowledge-based practices are still unevenly implemented and significant gaps remain between strategic intentions and everyday managerial practices (Laihonen & Kokko 2023; Laihonen et al. 2024). Considering these findings there is a need to examine how KHRM is implemented and developed across broader service systems and institutional environments, particularly from the perspective of HRM and leadership practices.

The purpose of this study is to explore and analyse how KHRM is implemented in Finnish public sector organisations through the lived experiences of employees involved in human resource management and leadership. By focusing on these experiences, the study seeks to develop an in-depth understanding of the current state and maturity of KHRM and examine how prepared these organisations are to move towards sustainable and strategic KHRM.

More specifically, this study aims to identify the organisational, managerial, cultural and competence-related factors that either enable or hinder the implementation and development of strategic and sustainable KHRM. By adopting an experience-based perspective, the study contributes to a nuanced understanding of how knowledge, data and analytics are currently utilized in human resource management practices and decision-making, and what conditions are required to strengthen KHRM as a long-term strategic capability. The research question of this study is:

Which factors enable and which factors hinder the implementation and development of sustainable and strategic KHRM in public sector organisations?

2 Theoretical Background

2.1 Sustainable human resources management

Sustainable HRM refers to an approach in which HRM is deliberately designed to generate simultaneous economic, social and human value in both the short and long term. Rather than focusing solely on efficiency or performance outcomes, sustainable HRM emphasizes the preservation and development of human and social resources over time, acknowledging employees as a critical and finite organisational asset (Kramar 2022.) From this perspective, sustainability in HRM is not an add-on to strategic HRM, but an extension that broadens its temporal horizon, stakeholder orientation and evaluative criteria.

A core assumption in sustainable HRM is that organisations face inherent tensions and trade-offs between competing goals, such as short-term productivity, employee well-being and long-term capability development. These tensions are particularly salient in a complex and resource-constrained context, such as the public sector, where organisations are expected to deliver high-quality services while maintaining legitimacy, fairness and social responsibility. Sustainable HRM literature highlights that such tensions cannot be resolved through isolated HR practices but requires a systemic and reflective approach to people management that recognizes paradoxes rather than attempting to eliminate them (Kramar 2014).

Within this framework, sustainable HRM places strong emphasis on fairness, organisational support, job satisfaction and employee commitment, which are viewed as both ethical imperatives and prerequisites for long-term organisational functioning (Cugueró-Escofet et al. 2019). Empirical research suggests that when employees perceive HR practices as fair, transparent and supportive, they are more likely to remain engaged, resilient and willing to contribute to organisational goals over time. These align with the objectives of sustainable HRM, as they support continuity, adaptability and the maintenance of human capital rather than short-term exploitation. (Pimenta et al. 2024)

Consequently, sustainable HRM highlights the need for decision-making that goes beyond financial indicators and incorporates information on employee-experiences, well-being and long-term capability development. From this perspective, the systematic use of information becomes prerequisite for sustainability-oriented HRM as it enables organisations to recognize the consequences of HR decisions across different time horizons and stakeholder groups (Aust et al. 2020). This creates conceptual link between sustainable HRM, strategic HRM and knowledge-based approaches of HRM.

2.2 Strategic human resources management

Strategic human resource management (SHRM) refers to the deliberate alignment of HRM practices, processes and decision-making with an organisation's long-term strategic objectives, based on the recognition of personnel as the key strategic resource. SHRM is proactive, systematic, and context specific. However, SHRM is not a universal or uniform model; its content and underlying mechanisms that are inherently context-dependent and vary across organisational and institutional settings (Knies et al. 2024; Lepak et al. 2006). In public organisations, SHRM primarily emphasises strategic coherence and a long-term development orientation rather than short-term efficiency gains. The central objective of SHRM in public sector value creation, reflecting the broader societal mission of public organisations. (Demo et al. 2024)

In the Finnish public sector, SHRM carries particular significance due to the operating environment characterised by short legal regulation, high societal expectations and persistent resource constraints. These conditions intensify the need for transparent, equitable and data-informed HR-practices that support both organisational legitimacy and sustainable performance (Jarlström, Saru & Vanhala 2018). While SHRM can positively influence employee motivation, attitudes and organisational performance in public organisations, these effects do not primarily arise through economic incentives of efficiency-oriented reward system, as is often the case in the private sector. Instead, public sector performance is shaped by multiple, simultaneous and

sometimes competing objectives, such as service quality, equality, legitimacy and the creation of public value. Consequently, within public sector SHRM, particular emphasis is placed on HR practices that strengthen employee capabilities and enhance opportunities for participation, engagement and professional development (Demo et al. 2024)

2.3 Knowledge-based human resources management

KHRM refers to an approach in which HRM is systematically grounded in the collection, integration and use of data, information and organisational knowledge to support HR-related decision-making. Rather than relying primarily on intuition or isolated metrics, KHRM emphasizes the transformation of HR data into actionable knowledge that informs strategic decisions regarding workforce development, performance, and long-term organisational capability (Marler & Boudreau 2017; Patil & Priya 2024).

In the context of SHRM, KHRM provides a mechanism through which HRM can be more closely aligned with organisational strategy. Prior research has shown that the strategic contribution of HRM depends not only on the presence of HR practices, but on the organisation's ability to generate and interpret information about how these practices influence employee skills, motivation and opportunities to contribute (Lepak et al. 2006; Marler & Bourdeau 2007). KHRM supports this alignment by enabling evidence-based assessments of workforce-related risks, trade-offs and development needs, there by strengthening the strategic role of HRM beyond administrative or reporting functions.

At the same time, KHRM is closely connected to sustainable HRM. Sustainable HRM highlights the need to balance short-term efficiency with employee well-being and long-term capability development, a task that requires visibility into the human consequences of HR decisions over time. Such balancing cannot be achieved without systematic and transparent use of information, as sustainability-related tensions often remain invisible in traditional performance indicators (Ehnert 2009; Kramar 2022). From this perspective, KHRM functions as an enabling infrastructure that makes the effects of HR practices on engagement, learning and retention visible and discussable, thereby supporting sustainable HRM.

An increasingly important aspect of KHRM concerns its level of maturity. Prior studies indicate that maturity is not determined by technology alone, but by the integration of data governance, analytical competence, managerial interpretation and decision-making routines (Jääskeläinen et al. 2022; Tenhovuori 2024). Higher levels of KHRM maturity are associated with more coherent HR-strategies, stronger alignment between HRM and organisational goals, and an increased ability to address sustainability-related trade-offs in a deliberate manner (Tenhovuori & Helander 2025).

Taken together, KHRM can be understood as a connecting concept that operationalizes both strategic and sustainable HRM. While strategic HRM defines the alignment of HRM with organisational objectives and sustainable HRM emphasizes long-term and multi-stakeholder outcomes, KHRM provides base through which these orientations can be enacted in practice.

4 Methodology

4.1 The collection of the data

The data for this study was collected in connection with a broader survey on the maturity level of human resource management in November–December 2025. The data was collected with an electronic questionnaire. A link to the electronic questionnaire was distributed in a network operating under the Ministry of Finance, which includes experts working in tasks related to the knowledge-based management of human resources in municipalities and wellbeing services counties. The members of the network were also given the opportunity to distribute a cover letter with a survey link in their own organisations. At the questionnaire, there were a total of 5 open-ended questions, and they were divided into themes related to: strategy and organisation of KHRM, resources and competence, HR data and information needs, HR metrics and HR-analytics and utilization of HR data. In addition to the five open-ended questions, the respondents were asked about the organisation in which they work and their job title. The background questions selected were used to ensure that the respondents belonged to the target audience. This minimised the collection of sensitive and identifiable information about the respondents. To maintain anonymity, the respondents were given ID numbers.

A total of 94 experts responded to the questionnaire. Of them, 33 answered to one or more open ended questions. In total, the data comprised 75 open-ended responses. Some of the answers did not answer the set research question and were therefore omitted from the final data. The final material consisted of 70 responses

submitted by 30 different respondents. The respondents worked in a total of 14 different organisations. The research material consisted of a total of four A4-sized pages of text, covering a total of 2133 words.

4.2 Analysis of the data

In the analysis phase of the data, the themes were abandoned. This was done in order to get an overall picture of the material and to better perceive the connections between the themes. The analysis proceeded in accordance with the topics of qualitative inductive content analysis. Inductive content analysis is recommended for the use in situations where there is little previous research on the phenomenon and the researcher wants to form an in-depth preliminary understanding of the topic (Lyhne et al. 2025).

The analysis of the data began with getting acquainted with the data. The material was carefully read through. The descriptive data was imported into Excel file. A coding template was developed based on the research question. The unit of analysis was a set of ideas that contained one or more sentences. The originality of the context was preserved in the reduction phase, and encoding the analysis units corresponding to the research question and exporting them to an Excel file made it possible to return to them. The reductions (n=163) were grouped into subcategories, main categories and further themes based on content homogeneity. Although the content analysis was carried out with special care, there is always the possibility of error (Boselie et al, 2005).

5 Results

The implementation of KHRM is primarily affected by factors related to HR reporting, analytics and indicators, the collection, availability and storage of HR data, organisational roles, structures and responsibilities, resources and competence, strategic steering and organisational culture, information systems and technical prerequisites, and the utilisation of data in decision-making. In each of the identified main categories, factors that promote and weaken management with HR knowledge were found. The results of the study are described by theme for the main categories and subcategories in Figure 1.

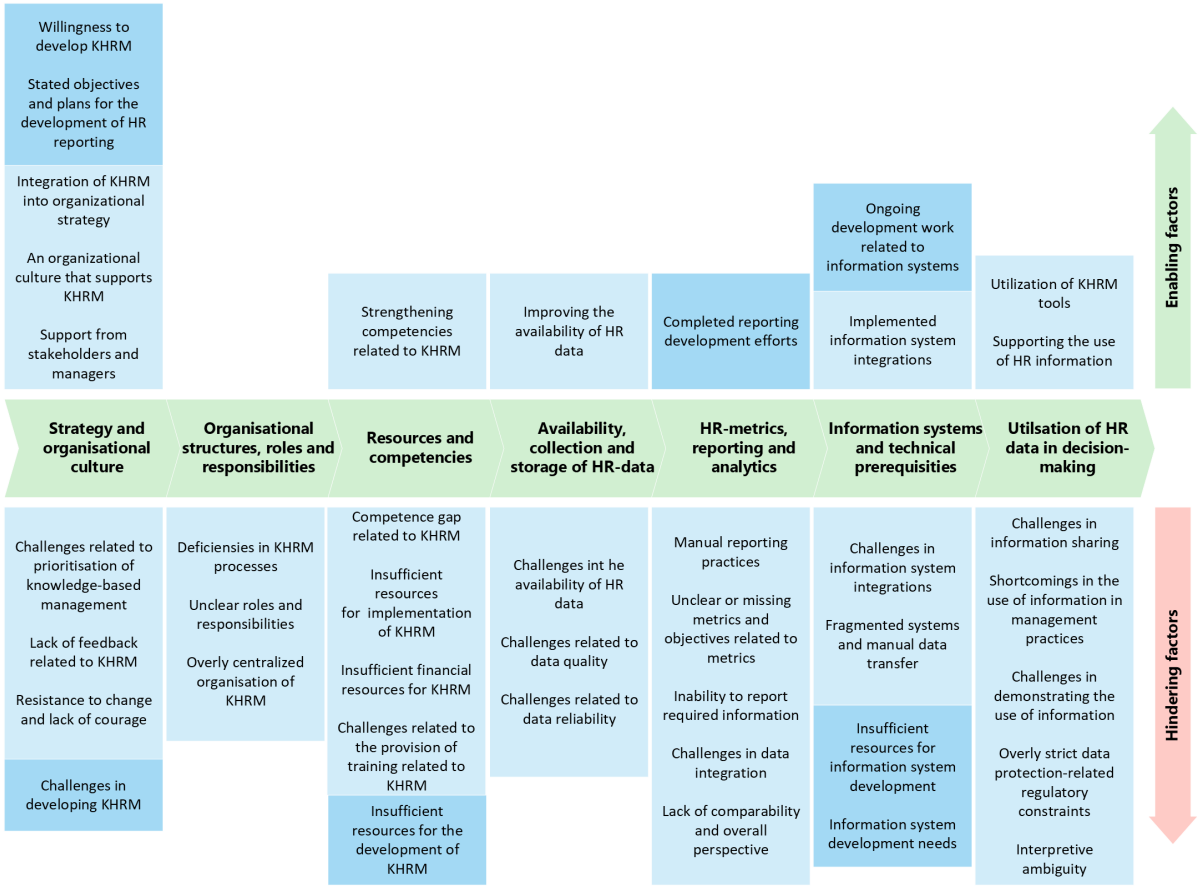


Figure 1: Themes, main categories and subcategories

5.1 Strategy and organisational culture

The development of KHRM is promoted by the goals and plans presented for the development of HR reporting (74, 25, 61, 89, 46) and the willingness to develop management based on HR information (88). The implementation of KHRM is promoted by the inclusion of KHRM in the organisation's strategy (25, 7, 55), the support received from senior management and supervisors (90) and the organisational culture that supports KHRM (88).

"KHRM is reflected in organisational strategy, and to some extent it is supported by organisational culture. There is a clear orientation towards monitoring HR indicators and using them to drive improvement."

The implementation of KHRM is weakened by resistance to change and lack of courage (43, 45, 61, 71, 74), prioritisation related to knowledge-based management (39, 43, 48, 68), challenges in developing KHRM (12, 27, 48, 78) as well as lack of feedback related to KHRM (78).

"Information is produced for the use of the field, but it remains unclear how this will be used in the field, or whether it will be used at all. We don't get feedback on whether the information produced is necessary, whether it is sufficient, or what we would like to see more of."

5.2 Organisational structures, roles and responsibilities

No factors promoting the implementation or development of KHRM were identified. On the other hand, deficiencies in KHRM processes (8, 68, 96), unclear roles and responsibilities (7, 8, 39, 40) and overly centralised organisation of KHRM were identified as factors hindering the implementation and development of KHRM (45).

"An overly centralised resource for knowledge-based management tends to forget the field where it should be used in everyday life."

5.3 Resources and competencies

The implementation of KHRM is constrained by insufficient organisational resources (10, 14, 19, 25, 27, 39, 43, 44, 88, 90), competence gaps related to KHRM knowledge (8, 35, 88), limited financial resources (46), and the challenges related to the provision of systematic training in KHRM (40, 88). Similarly, the further development of KHRM is hindered by the lack of resources for development of KHRM (40).

"The same human resources have many other tasks, which results in a lack of time to focus solely on the development of knowledge-based management."

5.4 Availability, collection and storage of HR-data

Improved availability of HR information strengthens the implementation of KHRM (25, 51). However, the effective implementation of KHRM is often constrained by challenges related to the reliability of HR data (21, 25, 88), limited availability of relevant HR data (27, 35) and challenges in the quality of HR data (68).

"More attention should be paid to the quality of data. It is directly proportional to the urgency of managerial work and the focus on HR matters."

5.5 HR-metrics, reporting and analytics

The implementation of KHRM is hindered by challenges related to HR reporting practices (34), the manual nature of HR-reporting processes (19, 39, 51), and the lack of clarity regarding HR metrics and the targets related to the metrics (12, 27, 35, 39, 40, 43, 51, 61, 96).

"There is a need for more shared discussion within organisation on KHRM, particularly regarding what is measured and for what purpose. At present, KHRM relies heavily on HR department, highlighting the need to make these practices more flexible and responsive to the information needs of the organisation. Strengthening organisation-wide engagement with HR-related data would support a more holistic and effective use of information in decision-making."

Additional challenges include difficulties in integrating HR data across systems (39, 40, 61), the lack of comparability and an overall holistic view of HR information (7, 88), as well as the need to further develop HR analytics capabilities (39, 49, 55) and predictive reporting practices (19, 39, 43, 49). At the same time, the development of KHRM is supported by prior development work already undertaken within the organisation (19).

"At present, there is a clear need to automate recurring HR reports. This would free up HR specialists' time for the meaningful use of HR-related information, such as examining patterns and conducting analysis, and would enable them to retrieve data from information systems for deeper, needs-based analysis when required."

5.6 Information systems and technical prerequisites

Completed integrations between information systems (25) and development work already carried out or currently underway in relation to HR information systems (25, 40, 46) support the implementation and further development of KHRM. However, implementation and further development of KHRM are hindered by insufficient resources allocated to information system development (35), challenges in system-to-system integrations (25, 68) as well as the fragmentation of information systems and the reliance on manual data transfer (27, 34, 40) and current development related needs related to HR information systems (12, 21, 27, 34, 40, 55).

"With the current systems, it is not possible to implement knowledge-based management as we would like without manual work."

5.7 Utilisation of HR data on decision-making

The implementation of KHRM is hindered by challenges related to knowledge sharing (12, 27, 40, 88, 90), data protection regulation (40, 88), the interpretative ambiguity of reports and analytics (88), insufficient integration of information use into managerial practices (43), and challenges related to verifying the actual use of information in decision-making (25, 40, 45, 51, 78).

"Senior management and those holding managerial positions utilise KM tools to support decision-making. However, the competence of middle management and frontline supervisors, as well as the implementation of decisions and the utilisation of information almost always require guidance from higher levels of management. Decisions may be made contrary to strong evidence or available information is simply not utilised. There is limited willingness to act differently from established practices. Frontline supervisors are particularly exposed to the influence of personnel. There is often an inability to justify decisions based on data, leading instead to reliance on established practices or the preferences of staff. Even when employees' views conflict with evidence derived from KM, decisions tend to follow previous practices or individual preferences."

In contrast, the implementation and further development of KHRM are supported by organisational efforts to actively promote the use of HR related information (25, 61) and the utilization of KHRM tools (45).

"At the moment, we are strengthening the foundation (availability of information, data quality, visualisation, up-to-date situational awareness for decision-making and support for the use of information) and resourcing in general to achieve the goals. At the end of 2026, the aim is to take big steps forward in this area in terms of the implementation of development measures."

6 Conclusions

This study set out to explore how KHRM is implemented and developed in Finnish public sector organisations through the experiences of employees involved in HRM and leadership. The study contributes empirical insights into practical realization of KHRM in context characterised by strong legal regulation, societal expectations and persistent resource constraints (Järlström, Saru & Vanhala, 2018).

This study identified several strategy-related factors that promote both the implementation and development of KHRM. Thus, the implementation and development of KHRM are strongly shaped by strategic and cultural factors. This supports the key assumptions of SHRM and sustainable HRM literature. In line with SHRM, the results highlight the importance of integrating KHRM into organisational strategies and the role of explicit stakeholder and manager support in creating legitimacy and direction for KHRM practices (Lepak et al. 2006; Knies et al. 2024). Furthermore, organisational cultures that are receptive to data-driven practices and oriented toward systematic development appear to function as enablers for KHRM, reinforcing the long-term development focus emphasised in sustainable HRM (Ehnert 2009; Kramar 2022). The results of this study reinforce prior research indicating that KM in public sector is inherently shaped by institutional complexity, where multiple logics, actors and governance mechanisms influence how information is produced, interpreted and utilised (Laihonen & Kokko 2023; Laihonen et al. 2024). In this context, the implementation of KHRM cannot be understood solely as a technical or organisational issue, but rather as a phenomenon deeply embedded in broader service systems, organisational culture and institutional conditions.

At the same time, the results reveal substantial barriers that hinder the maturation of KHRM. These barriers are not limited to technical challenges but extend to fragmented organisational structures, unclear roles and responsibilities, limited resources, competence gaps, and difficulties in embedding the use of HR information into everyday managerial practices. This finding is consistent with prior research on knowledge-based management and maturity models, which emphasise that maturity is determined by the alignment of data governance, analytical competence, organisational structures and decision-making routines rather than by

information systems alone (Jääskeläinen et al. 2022; Tenhovuori & Helander 2025). However, the study also points to persistent gap between strategic intent and operational reality, particularly in the utilisation, interpretation and verification of HR data in decision-making, which limits the realisation of strategic and sustainability-oriented HRM in practice.

None of the respondents of this study referred to the use of AI in HRM or knowledge-based management. This suggests that Finnish public sector organisations are largely situated at earlier stages of KHRM maturity, focusing on foundational capabilities such as data availability, data quality, information system integration and analytical competence. Several respondents explicitly highlighted current information system development needs or ongoing system renewal projects as factors hindering the effective implementation of KHRM, indicating that organisations are still engaged in strengthening basic technical and structural prerequisites. This observation supports the view that KHRM development should be understood as a staged, maturity-based process rather than as a technology-driven transformation (Marler & Boudreau 2017; Jääskeläinen et al 2022; Tenhovuori & Helander 2025). From a KHRM perspective, information systems and technical prerequisites form a critical foundation for effective use of HR data in decision-making, and limitations at this level constrain progress towards more advanced, analytics- or AI-enabled HRM practices (Tenhovuori 2024).

The findings related to resources further reinforce the previous literature on sustainable and strategic HRM. Consistent with sustainable HRM perspective, this study shows that insufficient organisational, financial and competence-related resources weaken the implementation and development of KHRM, despite widespread recognition of its strategic importance. This tension is particularly pronounced in the public sector, where service provision is fundamentally dependent on human resources and where personnel costs constitute a substantial share of organisational expenditure. Sustainable and strategic KHRM is expected to support the availability, capability and well-being of personnel to ensure the continuity, quality and effectiveness of public services. (Keva 2025) However, the results of this study indicate that limited resources, competing priorities and insufficient time allocated to systematic development undermine the realisation of these objectives in practice. This finding underscores a central paradox of sustainable HRM in public sector: while human resources are recognised as core enabler of public value creation, the resourcing of KHRM necessary to support this role remains insufficient.

7 Ethics Declaration

Ethics declaration was not required for this research. No personal data was collected in this study. Participation in the study was voluntary, and responding to the electronic questionnaire was considered as informed consent. In connection with the link to electronic questionnaire, the respondents were provided with cover letter. The respondents answered the electronic questionnaire anonymously. The research material is stored appropriately and carefully throughout the life cycle of the research project.

8 AI Declaration

Artificial intelligence tools (Microsoft Copilot) were used for language refinement only. All substantive content, analysis and conclusions were produced and validated by the authors.

References

Aust, I., Matthews, B. & Müller-Camen, M. (2020) Common good HRM: A paradigm shift in sustainable HRM? Human Resource Management Review, 30(3).

Boselie, P., Dietz, G. and Boon, C. (2005), “Commonalities and contradictions in HRM and performance research”, Human Resource Management Journal, Vol. 15 No. 3, pp. 67-94.

Brito, R. & Oliveira, L. (2022) Knowledge-based public management: implications for workforce planning and accountability. International Journal of Public Administration, 45(6), 473–486.

Chatterjee S., Shandhuri R., Vrontis D. & Sianchou E. (2021) Examining the dark side of human resource analytics: an empirial investigation using the privacy calculus approach. International Journal of Manpower. Vol 43, No. 1, 52-74

Cugueró-Escofet, N., Ficapal-Cusí, P. and Torrent-Sellens, J. (2019) Sustainable human resource management: How to create a knowledge sharing behavior through organizational justice, organizational support, satisfaction and commitment. Sustainability, 11(19), 5419.

Dahlbom P., Siikanen N., Sajasalo P. & Jarvenpää M. 2020. Big data and HR analytics in the digital era. Baltic Journal of Management. Vol 15 No 1, pp 120-138

Demo, G., Costa A., & Coura K. (2024) HRM practices in the public service: a measurement model. RAUSP Management Journal. Vol 59, No 1, 50-66. DOI: 10.1108/RAUSP-01-2023-0005

Ehnert, I. (2009) Sustainable Human Resource Management: A Conceptual and Exploratory Analysis. Berlin: Springer.

Erkkilä, T. (2020) Transparency in public administration. Oxford Research Encyclopedia of Politics. Oxford University Press. https://doi-org.libproxy.tuni.fi/10.1093/acrefore/9780190228637.013.1404

Jarlström, M., Saru, E. & Vanhala, S. (2018) Sustainable HRM with salience of stakeholders. Journal of Business Ethics, 152, 703–724.

Jääskeläinen, A., Sillanpää, V., Helander, N., Leskelä, R.-L., Haavisto, I., Laasonen, V. & Torkki, P., (2022). Designing a maturity model for analysing information and knowledge management in the public sector. VINE Journal of information and knowledge management systems 52, 120–140

Kauhanen, A. (2025) Julkisen sektorin palkanasetanta: Voidaanko Suomeen luoda vientivetoinen työmarkkinamalli? ETLA Raportit 163. Elinkeinoelämän tutkimuslaitos.

Keva (2025) Julkisen alan työhyvinvointi vuonna 2024. Keva tutkimuksia 2/2025. Julkisen alan työhyvinvointi vuonna 2024

Knies, E., Boselie, P., Gould-Willims, J. & Vandenabeele, W. (2024) Strategic human resource management and public sector performance: context matters. The International Journal of Human Resource Management, 35:14, 2432-2444, DOI: 10.1080/09585192.2017.1407088

Kramar, R. (2022) Sustainable human resource management: six defining characteristics. Asia Pacific Journal of Human Resources, 60(1), pp. 146–170.

Laihonen, H. & Kokko, P. (2023) 'Knowledge management and hybridity of institutional logics in public sector', Knowledge Management Research & Practice, 21(1), pp. 14–28. https://doi.org/10.1080/14778238.2020.1788429

Laihonen, H., Kork, A-A., Sinervo, L-M. (2024) Advancing public sector knowledge management: towards an understanding of knowledge formation in public administration. Knowledge Management Research & Practice, 22:3, 223-233. https://doi.org/10.1080/14778238.2023.2187719

Lepak, D., Liao, H., Chung, Y & Harden, E. (2006) A Conceptual Review of Human Resource Management Systemis in Strategic Human Resource Management. In: Martocchio, J.J. (ed.) Research in Personnel and Human Resources Management. Vol. 25. Bingley: Emerald Group Publishing Limited, pp. 217-271. https://doi.org/10.1016/S0742-7301(06)25006-0

Lyhne, C.N., Thisted, J. & Bjerrum, M. (2025) Qualitative content analysis – framing the analytical process of inductive content analysis to develop a sound study design. Quality & Quantity, 59, 5329–5349. https://doi.org/10.1007/s11135-025-02220-9

Marler J. & Bourdeau J. 2016. An evidence-based review of HR Analytics. The International Journal of Human Resource Management. Vol. 28, No. 1, 3-26.

Patil, B.S. & Priya, M.R. (2024). HR data analytics and evidence-based practice as a strategic business partner. VILAKSHAN – XIMB Journal of Management. Vol. 21 No. 1. pp. 114-125. https://doi.org/10.1108/XJM-07-2023-0148

Partanen, E., Raatikainen, P., Hellsten, P. & Myllärniemi, J. (2023) 'The maturity of knowledge-based management in Finnish central-government organizations', in Lindgren, I. et al. (eds.) Knowledge-Based Management in the Public Sector. Springer.

Pimenta, S., Durate, A.P. & Simoes, E. (2024) How socially responsible human resource management fosters work engagement: the role of perceived organizational support and affective organizational commitment. Social Responsibility Journal, 20(2), 326-343. https://doi.org/10.1108/SRJ-10-2022-0442

Tenhovuori, S. & Helander, N. (2025) Measuring the maturity of knowledge-based management of human resources. In: Proceedings of the 17th International Joint Conference on Knowledge Discovery, Knowledge engineering and Knowledge Management (IC3K 2025), Volume 2: KMIS. SciTePress, pp. 484-491. https://doi.org/10.5220/0013746100004000

Tenhovuori, S. (2024). The Maturity Level and Benefits of Human Resource Analytics in a Public Organization Providing Social and Health Services. In G. Morales Alonso & G. Schiuma (Eds.), Proceedings of the 19th International Forum on Knowledge Asset Dynamics (IFKAD 2024): Translating Knowledge into Innovation Dynamics (pp. 1081–1093). Madrid Spain: Institute of Knowledge Asset Management. ISBN 978-88-96687-17-8.

Vakkuri, J., Vinnari, E. & Vikstedt, E. (2023) 'Performance management in the Finnish government: Reflections and future directions from the perspectives of hybridity and sustainability', in Pekkola, E. et al. (eds.) Finnish Public Administration, Governance and Public Management. Springer.

Life Cycle Costing in Regenerative Systems: A Knowledge Integration Approach

Francesca Lerario, Achille Claudio Garavelli
(DMMM), Polytechnic University of Bari, Bari, Italy

Abstract

Post-extractive landscapes, such as abandoned quarries, represent complex socio-ecological systems in which environmental degradation, material flows, regulatory constraints, and local development dynamics intersect. Within this context, these environments can be interpreted as manifestations of a broader polycrisis, where multiple ecological, economic, and institutional pressures interact and reinforce one another. Regenerative approaches respond to this condition by moving beyond impact reduction toward restoration-oriented value creation processes that actively contribute to ecosystem and community recovery. At the same time, the implementation of regenerative systems requires lifecycle-oriented evaluation tools capable of supporting decision-making under conditions of uncertainty, long-term dynamics, and multi-stakeholder involvement. This paper develops an analytical framework, grounded in socio-ecological and regenerative system perspectives, for the application of Life Cycle Costing (LCC) to restoration systems based on circular materials and additive manufacturing. The proposed approach structures cost evaluation across lifecycle phases, including material processing, manufacturing, installation, monitoring, maintenance, and end-of-life, and considers alternative production configurations at different scales. The study highlights how lifecycle cost structures can support knowledge integration and stakeholder coordination, particularly in configurations where monitoring, maintenance, and data-driven activities play a central role. It also emphasizes the relevance of lifecycle-based analysis in capturing the temporal dimension of value creation and in supporting the alignment of technical and organizational decisions across system phases and improving transparency in complex, multi-actor restoration settings. Expected results suggest that LCC can enable the comparison of alternative restoration configurations and provide insights into the temporal distribution of activities and responsibilities across lifecycle phases. At the same time, the study acknowledges the limitations of LCC in capturing broader ecological and social value, pointing to the need for complementary evaluation approaches. The paper contributes to the literature on regenerative socio-ecological systems and lifecycle evaluation by proposing a structured framework that integrates economic analysis with knowledge-intensive and multi-stakeholder restoration processes.

Keywords: Regenerative approach; Life Cycle Costing; Knowledge Integration; Product-Service Systems; Polycrisis

Paper type: Academic Research

1 Introduction

Abandoned quarries are not merely degraded sites; rather, they can be interpreted as territorial manifestations of a slow-onset polycrisis: complex socio-ecological systems shaped by the interaction of environmental degradation, material flows, regulatory dynamics, and local development trajectories (Chenot, 2018; Legambiente, 2023; Legambiente, 2025). In recent years, such contexts have been increasingly interpreted within broader systemic perspectives that emphasize the interconnected nature of environmental, economic, and social challenges (Dirk Helbing, 2013; Lawrence et al., 2024). This framing highlights how post-extractive landscapes are embedded in multi-layered processes that require coordinated responses across multiple domains of expertise.

Within this perspective, quarry restoration represents not only an ecological or technical challenge, but also an organizational and coordination problem. Restoration processes typically involve a wide range of actors, including engineers, ecologists, material scientists, public authorities, and local stakeholders, whose knowledge bases, objectives, and time horizons may differ significantly. Research on socio-ecological systems emphasizes that outcomes in such contexts depend on the ability to manage interdependencies and coordinate heterogeneous forms of knowledge (Buckton et al., 2023). In practice, however, achieving such coordination remains difficult, particularly when restoration outcomes unfold over long timeframes and under conditions of uncertainty.

Recent contributions on regenerative business models (RBMs) further reinforce the need for systemic and coordinated approaches. This stream of literature moves beyond traditional sustainability and circular economy frameworks by focusing on value creation processes that aim to restore and enhance ecosystem health (Konietzko et al., 2023; Das & Bocken, 2024). Despite this conceptual advancement, empirical applications remain limited, especially in territorially embedded contexts such as post-extractive landscapes. Studies indicate

that organizations often face significant challenges in translating regenerative principles into operational practices, particularly when implementation requires coordination across multiple stakeholders and domains (Gualandris et al., 2023).

In parallel, Life Cycle Costing (LCC) has been widely adopted as a techno-economic tool for evaluating costs across the lifecycle of products and systems (European Commission, 2023; International Organization for Standardization, 2017). Within sustainability and circular economy contexts, LCC supports decision-making by enabling the comparison of alternative solutions based on their long-term cost implications (Kambanou & Lindahl, 2016; Kambanou & Sakao, 2020). However, the use of LCC remains primarily oriented toward economic evaluation, and its application in complex socio-ecological systems raises important limitations. In particular, LCC relies on assumptions regarding system boundaries, lifecycle duration, and future operating conditions, and it does not fully capture non-monetizable dimensions such as ecological functions or social value.

These limitations are especially relevant in quarry restoration contexts, where decision-making involves long-term ecological processes, uncertainty, and multiple interacting actors. At the same time, technological advances in additive manufacturing and circular materials, such as the use of PLA-based composites incorporating marble waste, demonstrate the technical feasibility of innovative restoration approaches (Fico et al., 2025; Romani et al., 2021). Nonetheless, existing studies tend to focus on material performance or process efficiency, often without integrating economic, ecological, and organizational dimensions within a unified analytical framework.

Against this background, this paper adopts Life Cycle Costing as an analytical tool to evaluate the economic performance of a regenerative restoration system for post-extractive environments. The study focuses on a restoration configuration based on additive manufacturing and circular materials, and aims to provide a structured comparison of lifecycle cost components across alternative system configurations. Accordingly, the paper addresses the following research question:

RQ: *How does Life Cycle Costing support the evaluation of economic performance in regenerative restoration systems for quarry sites?*

To answer this question, the paper evaluates: (1) evaluating the lifecycle costs of a regenerative restoration system based on additive manufacturing and circular materials; (2) comparing cost structures across different lifecycle phases; and (3) discussing the implications of cost distribution for decision-making processes, stakeholder coordination, and the understanding of dynamics in complex socio-ecological systems.

2 Theoretical Background

2.1 Regenerative socio-ecological systems and lifecycle-oriented value creation

While circular economy strategies primarily focus on resource efficiency and waste reduction, regenerative approaches emphasize the restoration and enhancement of ecological and social systems (Konietzko et al., 2023). Accordingly, regenerative business models aim not only to reduce negative impacts, but also to generate positive outcomes for ecosystems and communities (Das & Bocken, 2024). Recent literature conceptualizes regenerative systems as socio-ecological configurations characterized by the interdependence of ecological processes, social dynamics, economic activities, and institutional arrangements (Buckton et al., 2023). This perspective is particularly relevant for post-extractive landscapes, such as abandoned quarries, where environmental degradation is intertwined with land-use change, material flows, regulatory constraints, and local development trajectories.

This view is consistent with systemic perspectives on risk and polycrisis, which emphasize the interconnected nature of environmental and socio-economic challenges (Dirk Helbing, 2013; Lawrence et al., 2024). In this sense, quarry degradation can be understood as the outcome of interacting pressures across ecological, economic, and institutional domains, rather than as an isolated environmental issue. Empirical evidence further highlights the complexity of these contexts: reports on quarrying and circular economy in the construction sector point to persistent environmental and material management challenges (Legambiente, 2021, 2025), while studies on quarry degradation and pedogenesis identify specific ecological constraints affecting restoration processes (Chenot et al., 2018). Within such contexts, value creation cannot be reduced to a single intervention phase, but must be understood as distributed across the lifecycle of the system. Literature on Product-Service Systems and servitization supports this perspective by showing how value emerges from the integration of products, services, and ongoing activities over time (Kohtamäki et al., 2019). In digital and service-oriented configurations, value is increasingly generated through monitoring, maintenance, and data-driven interactions, rather than through the delivery of a physical artefact alone (Langley, 2022; Paiola et al., 2024). Accordingly,

regenerative restoration systems require lifecycle-oriented approaches capable of capturing both the temporal distribution of activities and the coordination of multiple actors.

2.2 Circular materials and additive manufacturing for regenerative restoration

Circular materials and additive manufacturing have been increasingly explored as approaches to reduce waste and enable new applications of secondary resources within production systems. In circular economy contexts, additive manufacturing allows the transformation of waste streams into new products through material reprocessing and localized production processes (Romani et al., 2021). In relation to post-extractive landscapes, recent studies have demonstrated the feasibility of using quarry waste as input material for additive manufacturing applications. Fico et al. (2025) show how marble waste can be processed into eco-friendly filament for 3D printing, supporting renaturalization processes in quarry environments. This approach connects material recovery with restoration interventions, linking circular economy strategies to regenerative practices.

Material performance represents a key constraint in the application of circular materials to restoration contexts. The behavior of biodegradable polymers, such as polylactic acid (PLA) and related composites, depends on thermal degradation, environmental conditions, and long-term stability (Ohkita & Lee, 2006; Urayama et al., 2002), which directly affects their suitability in environments exposed to natural processes.

However, existing contributions tend to focus on material and process performance, with limited attention to broader lifecycle implications. Economic costs, service requirements, and long-term operational activities are often not integrated into the analysis. This gap reinforces the relevance of lifecycle-oriented approaches, such as LCC, to complement technical assessments and support a more comprehensive evaluation of regenerative restoration solutions. These limitations highlight the need for lifecycle-oriented analytical approaches capable of integrating technical, economic, and operational dimensions, such as Life Cycle Costing.

2.3 Life Cycle Costing (LCC)

Life Cycle Costing (LCC) is a techno-economic method used to assess costs incurred throughout the lifecycle of a system, including acquisition, operation, maintenance, and end-of-life phases (European Commission, 2023; International Organization for Standardization, 2017). Its relevance lies in supporting decisions that extend beyond initial investment, particularly in contexts characterized by long-term cost implications. In Product-Service System contexts, LCC has been applied to analyze how costs are distributed across lifecycle phases, enabling the comparison of alternative configurations that combine products and services (Kambanou & Lindahl, 2016). Similarly, in circular economy applications, LCC supports the evaluation of different resource and design strategies by accounting for costs associated with use, maintenance, and end-of-life management (Kambanou & Sakao, 2020). Applied to regenerative restoration systems, LCC provides a structured framework to compare solutions that involve both material components and long-term service activities. This is particularly relevant when restoration interventions extend beyond installation and require ongoing monitoring, maintenance, and adaptation over time. In such contexts, system performance depends not only on technical and economic factors, but also on ecological processes that evolve over time and are not fully predictable (Chenot et al., 2018). Beyond its evaluative function, the application of LCC requires the integration of heterogeneous forms of knowledge. Lifecycle modelling depends on inputs related to ecological dynamics, material behavior, production processes, cost structures, and regulatory conditions, which are often distributed across different domains. According to the knowledge-based view, the integration of specialized knowledge is a key condition for effective decision-making in complex systems (Grant, 1996). Research on knowledge boundaries further highlights how differences in assumptions and interpretative frameworks can complicate the integration of such knowledge (Carlile, 2002, 2004). In this sense, developing an LCC model involves making explicit a set of assumptions and parameters that reflect contributions from multiple domains. As a result, LCC does not eliminate knowledge complexity, but requires it to be structured within a common analytical framework. Despite its usefulness, LCC remains limited to the economic dimension of system performance and does not fully capture ecological or social outcomes. It should therefore be understood as a partial evaluation tool that supports economic comparison within broader, multi-dimensional decision-making processes.

3 Methodology

3.1 Research design and approach

This study adopts a conceptual and analytical Life Cycle Costing (LCC) modelling approach aimed at structuring the evaluation of regenerative restoration systems at the design stage. The research is currently in progress, with data collection and model calibration ongoing.

The approach is consistent with LCC applications in early design phases, where cost modelling supports decision-making prior to system implementation (Asiedu & Gu, 1998; Kambanou & Lindahl, 2016; Rodríguez et al., 2020). In line with established LCC practices, the model considers the entire lifecycle of the system, from material acquisition to end-of-life, adopting a long-term perspective (Rebitzer & Hunkeler, 2003; Settanni et al., 2014).

Two alternative production configurations are defined to structure the analysis. Scenario A represents laboratory-scale production, while Scenario B represents a potential industrial-scale configuration. These scenarios provide a comparative framework for analysing how differences in scale and process organization may affect lifecycle cost structures.

3.2 Goal and scope definition

The functional unit is defined as **1m² of restored quarry over a six-year period**. This temporal boundary reflects the timeframe typically required for vegetation establishment and initial ecosystem recovery in degraded environments (Urayama et al., 2002; Ohkita et al., 2006). The system is conceptualized as a restoration configuration based on the deployment of biodegradable microhabitat devices designed to support substrate stabilization, moisture retention, and microclimate regulation. The number of devices per m² is defined as a modelling assumption based on spatial coverage and design parameters.

The system boundaries follow a cradle-to-grave perspective and include five lifecycle stages:

- materials extraction and processing;
- manufacturing;
- installation;
- use phase;
- end-of-life.

3.3 Life Cycle Costing model

The LCC model is structured to estimate cost components across lifecycle phases, enabling the comparison of alternative system configurations. Costs are organized into the categories defined in 4.2.

The total cost per device is expressed as:

$$C_{device} = C_{materials} + C_{preprocessing} + C_{manufacturing} + C_{installation} + C_{monitoring} + C_{EoL}$$

Costs are then normalized to the functional unit based on the number of devices per m².

The model adopts a deterministic structure, with explicit assumptions to ensure transparency and comparability across scenarios.

3.4 Data and assumptions

The model is based on a combination of literature data, technical parameters, and preliminary market-based estimates. Data collection is currently ongoing and aims to refine key cost components and operational assumptions.

Data inputs include:

- material and process parameters from additive manufacturing and circular materials literature (Fico et al., 2025)
- cost-related assumptions informed by sectoral data and market observations
- lifecycle assumptions related to system duration, maintenance requirements, and environmental conditions

Given the current stage of the research, the model should be interpreted as an evolving analytical framework. Further work will focus on data validation, parameter refinement, and empirical calibration.

4 Expected Results and Discussion

Based on the proposed Life Cycle Costing framework and the characteristics of regenerative restoration systems, it is expected that the distribution of costs across lifecycle phases will provide insights into how activities are organized and coordinated among stakeholders. In particular, cost structures are anticipated to reflect different coordination patterns depending on their temporal distribution. Higher upfront costs associated with design and production phases may indicate a concentration of coordination efforts in early project stages, while the presence of recurring costs related to maintenance, monitoring, and operational activities is expected to signal the need for ongoing coordination among multiple actors over time. These expected patterns are consistent with existing research on Product-Service Systems and servitization, which highlights how service-oriented configurations require continuous interaction and coordination across lifecycle phases (Kohtamäki et al., 2019; Kambanou & Lindahl, 2016). In the context of socio-ecological systems, such coordination is particularly critical due to the involvement of heterogeneous stakeholders and long-term system dynamics (Buckton et al., 2023).

4.1 Expected knowledge requirements and informational complexity

The development of the LCC model is expected to require the integration of multiple forms of knowledge, reflecting the inherent complexity of regenerative restoration systems. Specifically, it is anticipated that ecological knowledge (e.g., environmental dynamics), engineering and manufacturing knowledge (e.g., additive manufacturing processes), economic knowledge (e.g., cost structures and assumptions), and institutional knowledge (e.g., regulatory frameworks) will need to be combined within a single analytical framework. From a theoretical perspective, this aligns with the knowledge-based view of the firm, which emphasizes the importance of integrating distributed and specialized knowledge for effective decision-making (Grant, 1996). The LCC modelling process is therefore expected to make explicit the range of assumptions and inputs required, highlighting both the diversity of knowledge domains involved and the potential challenges in aligning them. Furthermore, it is expected that uncertainty associated with key parameters, such as lifecycle duration, maintenance needs, and environmental conditions, will play a significant role in shaping the results. This suggests that LCC outcomes should be interpreted as contingent on underlying assumptions and subject to revision as additional knowledge becomes available.

4.2 Expected implications for multi-capital value and decision-making

While the LCC analysis is expected to provide a structured assessment of economic performance, it is also anticipated that it will reveal important limitations in capturing the broader value generated by regenerative restoration systems. Ecological benefits, such as the restoration of ecosystem functions, and social impacts at the territorial level are unlikely to be fully reflected in monetary terms. This is consistent with literature on sustainability accounting, which highlights the challenges of representing multi-dimensional value through economic metrics alone (Bebbington et al., 2014). From this perspective, regenerative restoration processes can be understood as generating value across multiple forms of capital – natural, social, and produced – of which economic cost represents only one dimension. Accordingly, it is expected that decision-making in such contexts cannot rely exclusively on LCC results but should incorporate complementary qualitative and multi-dimensional evaluations. The role of LCC, in this sense, is anticipated to be that of providing a structured basis for economic comparison, while supporting broader discussions on value creation in complex socio-ecological systems.

5 Conclusions

This study addressed the need for lifecycle-oriented evaluation in regenerative restoration systems within post-extractive landscapes. By framing these contexts as complex socio-ecological systems shaped by polycrisis dynamics, the study highlighted the challenges of coordinating technical, ecological, and organizational dimensions over time.

An analytical framework based on Life Cycle Costing (LCC) was developed to structure cost evaluation across lifecycle phases in regenerative Product-Service Systems. The framework enables the comparison of alternative configurations and highlight. At the same time, the study acknowledges that LCC captures only the economic dimension of system performance and does not fully represent broader ecological and social value. This limitation points to the need for complementary approaches to support decision-making in complex socio-ecological systems.

6 Ethics Declaration

Ethical approval was not required for the study

7 AI Declaration

The authors declare that all intellectual contributions, including the conceptualization, methodology, and analytical framework presented in this paper, are their original work. The use of artificial intelligence tools (ChatGPT) was limited to supporting the drafting and refinement of the text. The authors take full responsibility for the content, interpretations, and conclusions of the manuscript.

References

Buckton et al. (2023). The regenerative lens: A conceptual framework for regenerative social-ecological systems. *One Earth*, 6(7), 824–842.

Caccamo, M., Pittino, D., & Tell, F. (2023). Boundary objects, knowledge integration, and innovation management: A systematic review. *Technovation*, 122, 102645.

Carlile, P. R. (2002). A pragmatic view of knowledge and boundaries: Boundary objects in new product development. *Organization Science*, 13(4), 442–455.

Carlile, P. R. (2004). Transferring, translating, and transforming: An integrative framework for managing knowledge across boundaries. *Organization Science*, 15(5), 555–568.

Chenot et al. (2018) "Impact of Quarry Exploitation and Disuse on Pedogenesis", *Catena*, Vol. 160, pp. 354–365

European Commission. (2023). Life-cycle costing. https://green-forum.ec.europa.eu

Fico et al. (2025) "From marble waste to eco-friendly filament for 3D printing to help renaturalization of quarries", *Sustainability*, vol. 17, no. 5.

Grant, R. M. (1996). Toward a knowledge-based theory of the firm. Strategic Management Journal, 17(S2), 109–122.

Gualandris et al. (2023) "Unchaining Supply Chains: Transformative Leaps Toward Regenerating Social-Ecological Systems", *Journal of Supply Chain Management*

Das, A. and Bocken, N. (2024) "Regenerative Business Strategies: A Database and Typology to Inspire Business Experimentation Towards Sustainability", *Sustainable Production and Consumption*, Vol. 49, pp. 529–544

Dirk Helbing, D. (2013). Globally networked risks and how to respond. Nature, 497, 51–59.

International Organization for Standardization. (2017). ISO 15686-5: Buildings and constructed assets, Service life planning, Part 5: Life-cycle costing.

Kambanou, M. L., & Lindahl, M. (2016). A literature review of life cycle costing in the product-service system context. *Procedia CIRP*, 47, 186–191.

Kambanou, M. L., & Sakao, T. (2020). Using life cycle costing to select circular measures. *Resources, Conservation and Recycling*, 155, 104650.

Kohtamäki, M., Parida, V., Oghazi, P., Gebauer, H., & Baines, T. (2019). Digital servitization business models in ecosystems. *Journal of Business Research*, 104, 380–392.

Konietzko, J., Das, A., & Bocken, N. (2023). Towards regenerative business models: A necessary shift? *Sustainable Production and Consumption*, 38, 372–388.

Langley, D. J. (2022). Digital product-service systems: The role of data in servitization. *Sustainability*, *14*(3), 1303.

Lawrence, M., et al. (2024). Global polycrisis: The causal mechanisms of crisis entanglement. *Global Sustainability*, 7, e6.

Legambiente (2021) Rapporto Cave. La Transizione dell'Economia Circolare nel Settore delle Costruzioni 2021, [online], Available at: https://www.legambiente.it/rapporti-e-osservatori/rapporto-cave-2021/

Legambiente (2025). La transizione dell'economia circolare nel settore delle costruzioni, [online], Available at: https://www.infobuildenergia.it/wp-content/uploads/2025/11/CAVE-report-finale-3.pdf

Ohkita, T.; Lee, S.H. (2006) "Thermal Degradation and Biodegradability of Poly (Lactic Acid)/Corn Starch Biocomposites. *J. Appl. Polym*." Sci., 100, 3009–3017.

Paiola, M., Grandinetti, R., Kowalkowski, C., & Rapaccini, M. (2024). Digital servitization strategies and business model innovation: The role of knowledge-intensive business services. *Journal of Engineering and Technology Management*, 74, 101846.

Romani, A., et al. (2021). Design, Materials, and Extrusion-Based Additive Manufacturing in Circular Economy Contexts: From Waste to New Products. *Sustainability*, 13(13), 7269.

Urayama, H.; Kanamori, T.; Kimura, Y. (2002) "Properties and Biodegradability of Polymer Blends of Poly(L-Lactide)s with Different Optical Purity of the Lactate Units". *Macromol.* Mater. Eng., 287, 116–121.

From Unstructured Inputs to Structured Knowledge: A Modular LLM-Based Approach for Enterprise Knowledge Extraction

Umair Ali Khan, Dmitry Kudryavtsev, Janne Kauttonen, Akseli Leskinen, Jukka Remes
Haaga-Helia University of Applied Sciences, Helsinki, Finland

Abstract

Enterprises increasingly use generative AI to convert unstructured organizational knowledge into structured outputs that can support reporting, decision-making, analytics, and process automation. However, structured knowledge extraction in real enterprise settings remains difficult because input data often appears in fragmented forms such as speech transcripts, scanned documents, complex forms, and layout-rich business records. In addition, each use case usually requires its own fields, constraints, terminology, and output format, making tailored extraction pipelines costly to build and difficult to maintain.

This paper presents a modular, large language model (LLM)-based knowledge extraction component developed as part of the GAIK toolkit - a generative AI toolkit for supporting organizational knowledge processes. The component enables users to define extraction requirements in plain language, automatically generate a structured schema, and use that schema for the extraction of knowledge from parsed documents or transcribed speech using LLM. The approach is designed to reduce repetitive engineering work while improving consistency, reusability, and adaptability across different enterprise use cases.

The study follows a design science research approach. The artifact was designed, implemented, demonstrated, and evaluated through industry-oriented use cases, with a particular focus on incident reporting and construction site diary creation. The incident reporting case demonstrates how spoken safety observations can be converted into structured reports using a reusable audio-to-structured-data module that combines transcription, transcript enhancement, schema generation, and extraction. A pilot evaluation with 15 incident-reporting audio samples provides preliminary evidence of technical feasibility and shows how schema persistence, human review, and configurable semantic similarity thresholds can support adaptation when business requirements or field definitions change, while maintaining reusable knowledge extraction workflows across organizational contexts.

The paper presents a practical, reusable design for enterprise knowledge extraction using generative AI. It shows how plain-language requirements, dynamic schema generation, and schema-conformant extraction can be combined into configurable building blocks for knowledge capture.

Keywords: structured knowledge extraction, knowledge capture, dynamic schema generation, large language models, workflow automation, knowledge management

Paper type: Practical Paper

1 Introduction

Knowledge management focuses on how organizations capture, store, share, and reuse knowledge to support decisions, operations, and learning. Much of this knowledge is embedded in unstructured or semi-structured materials, such as meeting recordings, field notes, purchase orders, requests for proposals, inspection reports, and safety observations. When such content remains scattered across documents, systems, and individual employees, it is difficult to reuse, analyze, or integrate into business processes. Structured knowledge extraction addresses this problem by transforming organizational content into machine-readable records that support search, analytics, reporting, automation, and governance.

Generative AI, and especially LLMs, create new opportunities for supporting knowledge processes. Companies see especially high perceived usefulness of AI for knowledge acquisition and documentation, including eliciting personal knowledge from experts and extracting knowledge from discussions, meetings, and conversations. Recent work shows that LLMs can support a wide range of extraction tasks across domains (Ateia et al., 2025a). From a knowledge management perspective, this reflects a broader human-AI partnership, where humans define meaning, goals, context, and accountability, while AI helps scale knowledge capture, structuring, and reuse (Jarrahi et al., 2023).

However, applying LLMs to structured knowledge extraction in enterprises remains challenging. First, reliable extraction requires stable schemas with clear field names, data types, constraints, and optional fields. Without such schemas, LLM outputs may vary across runs and become difficult to validate or integrate. Second, enterprise inputs often include scanned documents, tables, forms, images, and speech transcripts. If these are

not parsed correctly, extraction quality decreases. Third, organizations need many extraction tasks across departments, each with its own terminology, fields, and reporting requirements. Building a separate solution for every use case is costly, hard to maintain, and poorly suited to SMEs with limited technical capacity.

These challenges are closely connected to the knowledge management idea of codification. Codification strategies aim to make knowledge explicit, reusable, and accessible through structured repositories and systems (Alavi and Leidner, 2001). In this paper, we argue that LLM-based structured extraction can support codification by converting tacit, fragmented, or document-bound knowledge into consistent records. This also raises a broader question: whether such systems enable new forms of organizational memory, understood as stored information from an organization's history that can inform present decisions (Walsh and Ungson, 1991). If spoken observations, field notes, reports, and business documents can be continuously converted into schema-conformant records, generative AI may expand how organizations capture and reuse memory from everyday work.

This paper addresses these issues through the design and evaluation of a modular LLM-based knowledge extraction component developed as part of the Generative AI Enhanced Knowledge Management (GAIK) toolkit (GAIK Project, 2026a). The component supports knowledge capture and codification by transforming fragmented workplace observations, speech records, and documents into reusable structured knowledge. The work is relevant to human-AI hybrid knowledge ecosystems, digital transformation of knowledge management practices, and generative AI for knowledge management.

GAIK is an open-source, business-oriented toolkit that organizes reusable GenAI building blocks around key knowledge processes, including knowledge capture, knowledge access, and knowledge synthesis (Kudryavtsev et al., 2025, 2024). The component presented in this paper focuses on knowledge capture. It allows users to describe extraction requirements in plain language, automatically converts them into type-safe schemas, and uses the generated schemas to guide schema-conformant extraction from parsed documents or speech transcripts.

The research is grounded in our experience from AI consultancy and industry-oriented development work, where structured knowledge extraction repeatedly appears as a bottleneck in early AI adoption. Common needs include extracting fields from purchase orders, invoices, and contracts; converting voice notes into structured reports; capturing observations from field work; and preparing consistent inputs for reporting, compliance review, or proposal generation. These recurring needs motivated reusable extraction building blocks rather than isolated use-case-specific scripts.

The paper is guided by the following research questions:

- *RQ1: How can plain-language extraction requirements be transformed into reusable schemas for structured enterprise knowledge extraction?*
- *RQ2: How can schema-guided LLM extraction be implemented as a modular building block for knowledge capture across documents and speech transcripts?*
- *RQ3: Can this approach support organizational memory by converting tacit and/or unstructured inputs into validated and reusable knowledge records?*

To answer these questions, the paper follows a design science research approach. We present the design of the extraction component, its integration into reusable software modules, and its demonstration through enterprise use cases, focusing on incident reporting and construction site diary creation. We also report an objective evaluation of the incident reporting use case, where spoken safety observations are transformed into structured reports through an audio-to-structured-data workflow.

2 Related Work

2.1 Academic work

(Xu et al., 2024a) conclude that LLMs can handle multiple extraction tasks in a unified format, with strong performance in zero-shot and few-shot settings. This generalization capacity is directly relevant to enterprise knowledge extraction, where extraction targets vary widely across departments and use cases, and annotated training data is rarely available at scale.

Within the same space, Ateia et al. (2025b) apply LLM for in-context learning to extract key concepts from scientific publications. Their approach requires few or no labelled examples to configure extraction for a new target field. This finding strengthens the case for plain-language, schema-driven extraction as a viable model for enterprise settings where requirements change frequently, and technical configuration capacity is limited.

Structured extraction from domain-specific text has also gained traction. Dagdelen et al. (2024a) demonstrate that LLMs can be fine-tuned for joint named entity recognition and relation extraction in materials

chemistry, producing structured records. Their work shows that LLMs can be directed to extract records consistently across sentences and paragraphs, and that the output format itself can be controlled through prompt design and schema specification.

A recurring challenge in LLM-based enterprise extraction is the accurate parsing of business documents. Zhang et al. (2024) note that VLM-based approaches offer stronger generalization, which is consistent with a recent study (Zhang et al., 2026). This forms the basis for one of our custom parsers (a multimodal parser) used in the software modules, where extraction is followed by parsing.

Schema-guided LLM-based extraction has emerged as a reliable method in production settings. When an LLM is guided to produce output conforming to an explicitly defined schema, the result is more consistent, easier to validate, and more directly integrable with downstream systems than free-form generation. Dagdelen et al. 2024b) and Xu et al., 2024b) both identify well-defined output schemas as critical for reliable, reproducible extraction, especially in settings where outputs must feed into reporting workflows, compliance systems, or databases.

The challenge that follows naturally from schema-guided extraction is schema generation itself. In enterprise settings, extraction requirements are often initially held as tacit knowledge, where a subject-matter expert knows what fields a report should contain. However, the mapping from that knowledge to a formal, machine-enforceable schema is a non-trivial step. If users must specify schemas manually in technical formats, the barrier to adoption rises significantly for organizations without strong AI or software engineering capability.

2.2 Industrial solutions

Many cloud platforms provide knowledge extraction services. Azure AI Document Intelligence (Microsoft Azure, 2026) extracts text, tables, document structure, and key-value pairs, and also supports custom models for structured, semi-structured, and unstructured documents. Google Document AI(Google Cloud, 2026) provides processors for parsing documents and extracting structured fields. Amazon Textract (Amazon Web Services, 2026) extracts text, handwriting, layout elements, forms, tables, and bounding-box information from scanned documents and images.

Current LLM APIs and developer frameworks increasingly support structured output generation. OpenAI Structured Outputs (OpenAI, 2026) are designed to make model responses adhere to a given schema. Azure OpenAI similarly describes structured outputs as a way to make a model follow a schema definition. Similarly, Anthropic(Anthropic, 2026) and LangChain(LangChain, 2026) also support structured outputs with a schema in various formats. These methods reduce the need to parse free-form LLM responses and make generated outputs easier to validate and integrate.

However, these industrial solutions do not fully address the problem studied in this paper. Document AI platforms are strong at document processing, but they are usually organized around predefined processors, custom models, or vendor-specific workflows. LLM structured-output APIs provide schema adherence, but they do not by themselves solve the broader enterprise problem of translating plain-language business requirements into reusable extraction schemas, selecting appropriate parsing or transcription routes, and embedding extraction into knowledge-management workflows. The contribution of this paper is therefore not another standalone extraction model, but a reusable design for configurable knowledge extraction as part of a broader GenAI toolkit for organizational knowledge capture.

3 Research Methodology

The research paper follows a design science research methodology (Peffers et al., 2007). Knowledge extraction components are part of the GAIK toolkit, and the overall methodology for toolkit development is presented in Kudryavtsev et al. (2026). Design and development of the required components were done in cooperation with companies in several iterations with a feedback loop. The knowledge extraction problem was identified through our AI consultancy in the Finnish AI Region project (Finnish AI Region, 2026) and research activities with organizations exploring generative AI adoption, where structured knowledge extraction appeared as a recurring problem. Based on this problem diagnosis, three design requirements were defined for the knowledge extraction solution: i) the artifact should provide business-level support, ii) it should ensure technical reuse, and iii) it should be designed for SMEs and organizations with low AI maturity, where users may not have the expertise to manually define schemas, validation rules, or complex LLM pipelines. The knowledge extraction solution design and development results are presented in Section 4. Demonstration and evaluation of the suggested artifact are presented in Section 5.

4 Knowledge Extraction Approach and Reusable Components

The resulting artifact is a modular knowledge extraction software component in GAIK's toolkit (GAIK extractor (GAIK Project, 2026b). The artifact consists of four main methodological steps (see Figure 1).

1. **Requirement parsing:** the user's natural-language specification is converted into an intermediate requirement representation. This representation defines the target fields, meanings, expected data types, required or optional status, and constraints such as enumerations or patterns.
2. **Structure detection:** the component determines whether the expected output is a flat record or a nested structure. A flat structure may involve extracting invoice-level fields such as invoice number, sender, receiver, date, subtotal, tax, and total. A nested structure may involve extracting repeated line items from a purchase order and linking each item to information from an accompanying bill of materials.
3. **Input preparation:** for documents, the input is first converted into a textual representation (called parsing). Simple documents are handled through standard parsers, while scanned, layout-rich, or visually complex documents are parsed with custom vision-based or multimodal parsers. This design is important because parsing quality directly affects extraction quality. If tables, charts, figures, or layout cues are incorrectly represented, the downstream extractor may receive incomplete or misleading input. Similarly, audio/video input is first converted into a textual format (called transcription).
4. **Schema generation and extraction:** extraction requirements are compiled into runtime validation models, enabling type checking, strict field control, and consistent output structures. The extractor then applies schema-guided extraction to the parsed input and returns schema-conformant structured records. This design follows recent work on LLM-based structured extraction, where schema constraints and structured outputs are used to reduce output drift and improve downstream integration (Dagdelen et al., 2024b; Xu et al., 2024b).

The artifact was operationalized through two reusable software modules in the GAIK toolkit. The first module, documents-to-structured-data, combines a document parser with the schema generator and extractor. It parses the input with a selected parser, generates or reuses an extraction schema from user requirements, and returns extracted fields, the generated schema, and the corresponding requirements. It also supports schema persistence, allowing organizations to generate a schema once and reuse it for repeated extraction tasks.

The second module, audio-to-structured-data, combines transcription with the same requirement-to-schema and schema-guided extraction pipeline. It accepts audio or video input, produces a raw transcript, applies transcript enhancement using our 2-pass transcript enhancement software component, and then passes the transcript to the extractor. The module returns extracted fields, a generated schema, and extraction requirements. The current transcriber component supports configurable transcription backends, including Whisper, GPT-4o-transcribe, and language-specific Whisper models. Figure 2 shows the two reusable knowledge extraction software modules using the knowledge extraction pipeline.

In the evaluation design, three validation dimensions were defined. First, functional validity assesses whether the component can transform plain-language requirements into schemas and produce schema-conformant outputs. Second, reusability and configurability assess whether the same component can support different extraction tasks by changing requirements, parser choices, or module configuration rather than rewriting the pipeline. Third, output quality assesses whether extracted fields match human-annotated ground truth. Exact fields, such as dates, names, and prices, are compared through normalized string or value matching, while semantic fields are compared through embedding-based cosine similarity. The performance metrics include precision, recall, F1 score, exact match rate, and semantic match rate.

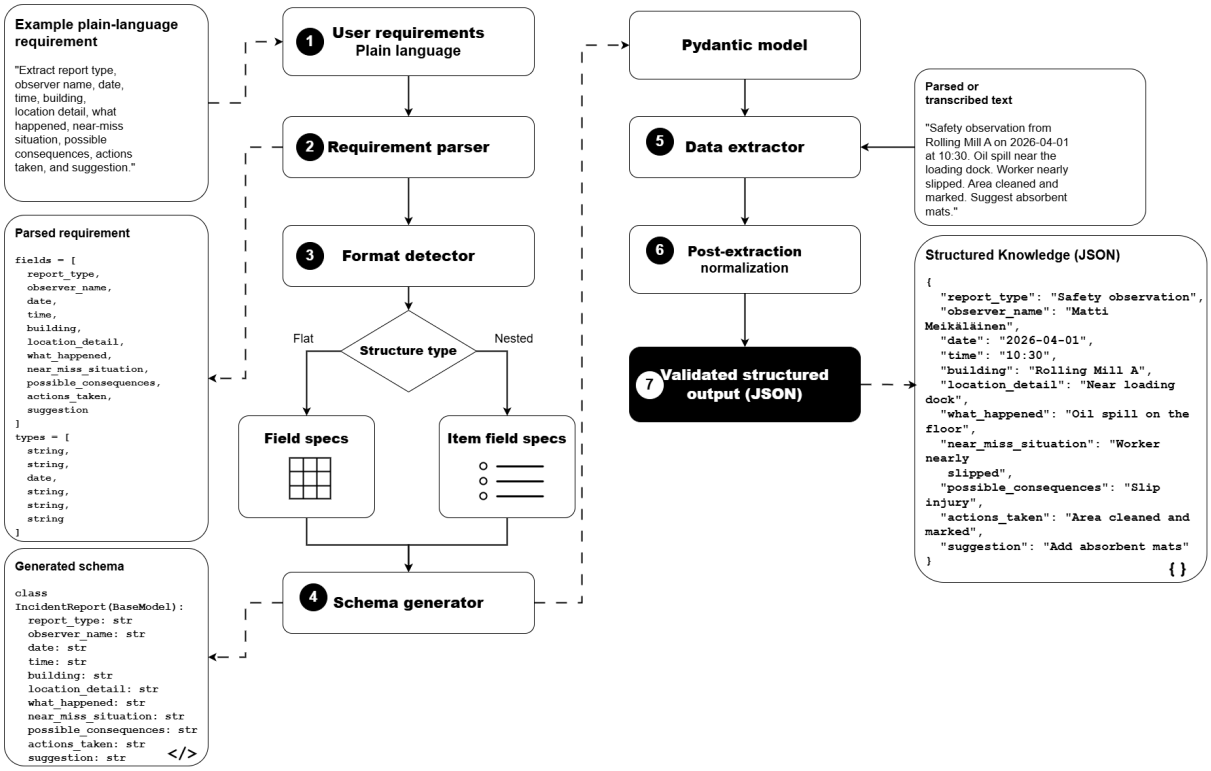


Figure 1: The main steps of the reusable knowledge extraction software component

5 Demonstration and Evaluation

5.1 Overview of use cases

The current implementation demonstrates end-to-end feasibility and reusability for several use cases, including incident report and safety observation in factories, construction site daily diary creation, construction site inspection report generation, and extracting data from purchase orders and bill of materials. The present focus is functional validation, modularity, and use case coverage rather than large-scale benchmarking.

Table 1: Two use cases for demonstration and evaluation

Use case name	Company	Problem	Solution
Incident reporting and safety observation	Manufacturing company	Reporting safety incidents is slow due to manual form-filling	Voice and image-based reporting with automatic structuring into reports
Construction site diary creation	Construction company	Difficult and time-consuming to document daily site activities	Converts voice notes, calls, and photos into structured daily diaries

5.2 Incident reporting and safety observations as one of the application areas for the extraction component

The incident reporting use case demonstrates how the knowledge extraction component can support structured knowledge capture from spoken workplace observations. In the partner company, employees are expected to report safety observations, near misses, and incidents such as spills, water leakage, broken equipment, or other hazards. The existing process requires employees to access a computer and complete a web form in the company's Impact system, which slows down reporting, especially when observations are made far from a workstation.

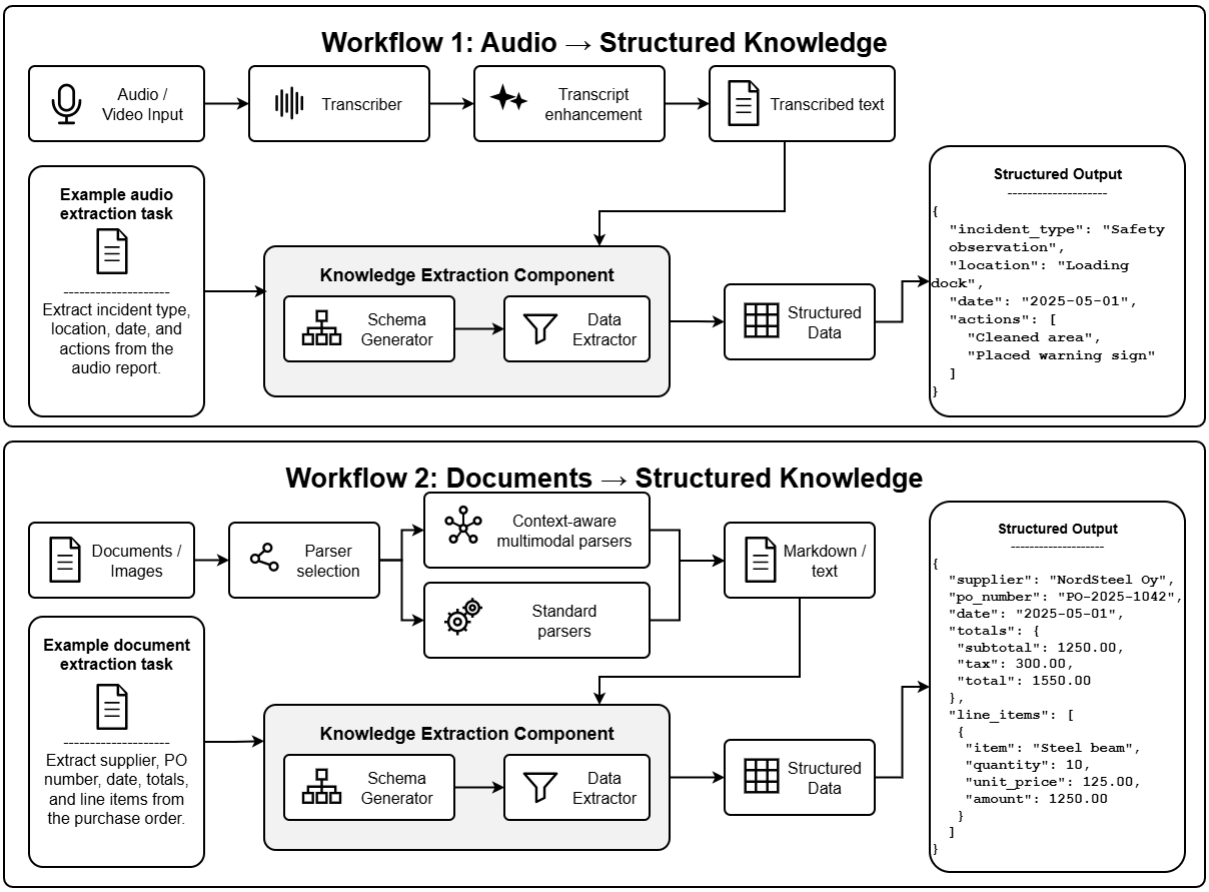


Figure 2: Depiction of the reusable knowledge extraction software modules using the knowledge extraction component

The designed AI-supported workflow enables employees to report incidents through voice input on a mobile device, with optional images when visual evidence is useful. The audio description is first transcribed by selecting a suitable transcription model, the transcript is enhanced using our 2-pass LLM-based enhancement method, and the extractor component then converts the description into the required structured fields of the incident report. The user reviews the extracted fields and modifies them as required. Subsequently, the generated report (raw and PDF converted) is transferred to the company's existing reporting system.

AI solution customization does not require technical skill. The schema generator helps convert natural-language requirements into a schema (e.g., an incident-reporting or safety-observation template). Once a schema is generated, it can be reused for the same extraction task. If, however, the task changes or some of the fields in the extraction task are modified, a new schema is generated and persisted. Figure 3 shows the solution implementation for incident reporting and safety observations.

5.3 Solution evaluation for incident reporting and safety observations

The evaluation of the incident reporting workflow mainly focused on the quality of the system output. Two other evaluation streams, system use and adoption evaluation, and business value evaluation, are in progress and will be conducted with the partner company after further piloting.

For the pilot evaluation, the partner company generated a dataset of 15 incident-reporting audio samples. Each audio described one incident, near miss, or safety observation. A blank structured schema (JSON) was prepared with company representatives to define the required report fields. This pilot primarily evaluated schema-guided extraction quality; systematic evaluation of automatic schema-generation accuracy remains future work. Company representatives listened to each audio sample and filled in the corresponding ground-truth values using this schema. This resulted in one ground-truth JSON file for each audio recording. Subsequently, the incident reporting AI workflow was applied to the same audio samples. The workflow transcribed the audio, enhanced the transcript, and used the approved schema and extractor to produce predicted JSON outputs.

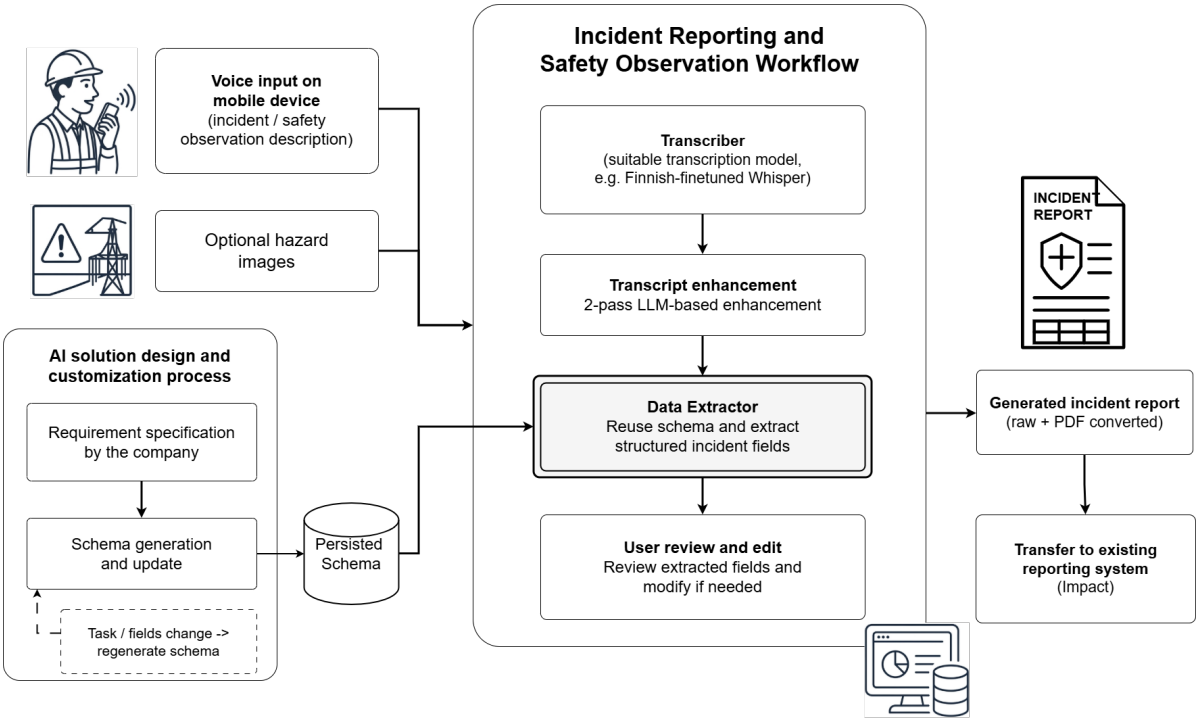


Figure 3: Solution implementation for incident reporting and Safety observations

The same ground-truth and predicted JSON files were then used for two types of system-output evaluation. In the subjective evaluation, members of the research team and company officials compared the predicted JSON outputs with the ground truth and used the feedback to refine the extraction task and prompt. In the objective evaluation, the predicted JSON outputs were compared against the ground-truth JSON files using field-level metrics.

The evaluation schema included two types of fields: i) those requiring exact agreement, such as report type, observer name, observer organization, date, time, etc., and ii) descriptive fields that may be correct even when expressed using slightly different wording, such as location detail, what happened, near-miss situation, possible consequences, actions taken, and suggestions. Exact fields were compared using normalized string or value matching. Semantic fields were compared using embedding-based cosine similarity with a threshold of 0.50. Fields scoring below 0.50 were classified as mismatches. This threshold was selected after analysis of several ground-truth and predicted pairs.

In comparison, a positive case means that a field has a non-empty ground-truth value. Therefore, a True Positive (TP) is counted only when the system correctly extracts a required non-empty value. If both the ground truth and prediction are empty, the comparison is counted as a True Negative (TN), not as a true positive, because no positive field value was present. If the ground truth is empty but the system produces a value, the comparison is counted as a False Positive (FP). If the ground truth contains a value but the system leaves the field empty, the comparison is counted as a False Negative (FN). If both contain values, but the values do not match, the comparison is counted as both a false positive and a false negative. This represents a strict field-level evaluation design where the system produced an incorrect value and, at the same time, failed to produce the correct ground-truth value.

Figure 4 shows an example of a field-level comparison of ground truth and AI-extracted JSONs. The objective evaluation was implemented using a Python script that iterated over matched ground-truth and prediction JSON files. Exact fields were normalized by lowercasing, collapsing whitespace, and stripping punctuation. Semantic fields were compared using OpenAI text-embedding-3-large embeddings and cosine similarity. The script produced field-level and aggregate results for the incident reporting extraction task.

The objective evaluation used five metrics: precision, recall, F1 score, exact match rate, and semantic match rate. Precision indicates the proportion of system-filled fields that were correct, while recall indicates the proportion of expected fields that the system successfully captured. The F1 score provides a combined view of precision and recall. The evaluation results are shown in Table 1.

The system captured most of the required information, with 90.00% precision, 87.10% recall, and an F1 score of 88.52%. This indicates that most system-filled fields were correct and that most fields that should have been captured were successfully extracted. The exact match rate of 90.67% shows strong performance on fields that

require normalized string or value matching. The semantic match rate of 87.78% shows strong performance on descriptive fields with the selected threshold of > 0.50.

The semantic match rate depends strongly on the selected similarity threshold. In this evaluation, a cosine similarity threshold of > 0.50 was used to classify semantic fields as matches. A higher threshold would be stricter and may reduce the number of accepted matches, especially when the system produces a correct but differently worded description. The threshold should therefore be selected based on the business use case. For low-risk applications, a lower threshold may be acceptable to capture more useful information. For safety-critical or compliance-related reporting, a stricter threshold may be preferable.

Table 2. Objective evaluation results for the incident-reporting workflow

Metric	Result	Interpretation
Precision	90.00%	Of all fields filled by the AI system, about 90% were correct.
Recall	87.10%	Of all fields that should have been filled, about 87.52% were captured.
F1 score	88.52%	Overall balance between precision and recall.
Exact match rate	90.67%	Strong performance on exact fields such as report type, observer name, date, and categorical values.
Semantic match rate (threshold > 0.50)	87.78%	Strong performance on descriptive fields where wording can vary.

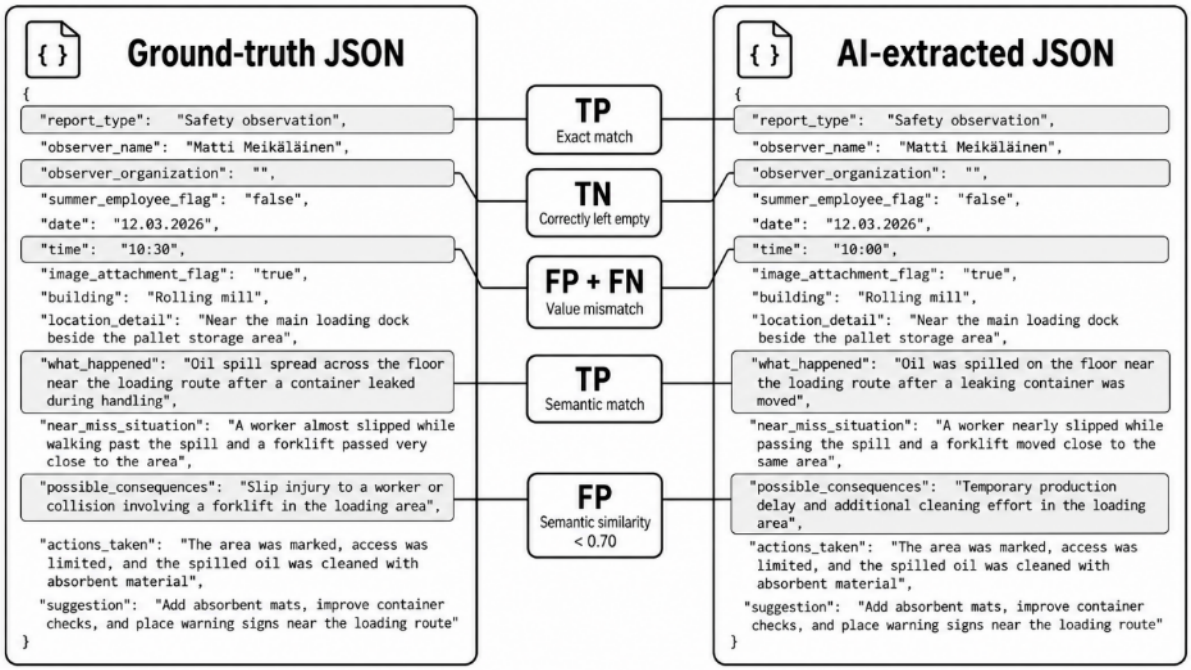


Figure 4: An example of a field-level comparison of ground truth and AI-extracted JSONs

5.4 Generalized publicly available and customisable knowledge extraction solutions

In addition to company-specific demonstrations and evaluations, we developed generalized and public demo solutions to show how the knowledge extraction component can be reused and customized across contexts. These demos are available through the GAIK demo portal: https://gaik-demo.2.rahtiapp.fi/.

The demo portal includes the standalone Knowledge Extraction Component, which allows users to define extraction requirements and test structured extraction directly. It also includes two reusable software modules that use the knowledge extraction component as a primary building block. The first module, Audio → Structured Data, transcribes audio and extracts structured data from the resulting transcript. The second module, Document → Structured Data, parses documents or images and extracts structured data from the parsed content.

Beyond these component-level and module-level demos, the portal also provides generalized business-level solutions built around knowledge extraction. The Incident Reporting demo allows users to report incidents through voice or text, after which AI extracts the required structured fields. The Construction Site Diary demo supports daily site documentation by allowing users to record activities through voice or text and convert them into structured diary entries.

These demos show that the proposed artifact is not limited to one partner-specific implementation. Instead, the same extraction logic can be exposed as a reusable component, embedded in software modules, and adapted into business-facing applications for different organizational needs.

6 Limitations

This study has some limitations. The evaluation is based on a small pilot dataset of 15 incident-reporting audio samples and one primary evaluated use case. The results, therefore, provide preliminary evidence of technical feasibility rather than generalizable performance across organizations, languages, document types, or reporting tasks. The evaluation also measures the end-to-end workflow and does not isolate errors from transcription, transcript enhancement, schema design, or extraction. In addition, the study does not include a baseline comparison or assess long-term adoption, business value, or actual reuse of extracted records as organizational memory. Future work should address these limitations through larger multi-case evaluations, baseline and ablation studies, and user-centered assessment of adoption, trust, and knowledge reuse.

7 Conclusion

This paper presented a modular LLM-based knowledge extraction component developed as part of the GAIK toolkit. The component addresses organizations' recurring problem of extracting structured knowledge from fragmented data. The proposed artifact allows users to describe extraction requirements in plain language, generate a reusable schema, and apply schema-guided extraction to documents or speech transcripts.

The study followed a design science research approach. The artifact was designed in response to practical needs identified through AI consultancy and industry collaboration, implemented as reusable software components, and demonstrated through company-specific and generalized public demos. The two reusable modules, audio-to-structured-data and document-to-structured-data, show how the same knowledge extraction component can be combined with transcription, transcript enhancement, document parsing, and multimodal parsing workflows.

The incident reporting evaluation provides initial evidence of technical feasibility and practical usefulness. From a knowledge management perspective, the artifact supports knowledge capture and codification by transforming fragmented organizational inputs into schema-conformant structured records. It can help make spoken observations, field notes, and documents easier to store, search, and reuse. However, this study does not evaluate long-term reuse or decision impact. Therefore, the approach should be understood as a human-AI knowledge-capture process rather than a fully autonomous replacement for human judgment in which AI performs the first structuring step, while humans review, correct, and approve outputs before integration into business systems.

Future work will expand testing across more companies, languages, document types, and extraction tasks; improve prompt and schema handling for sparse fields; explore alternative semantic thresholds; and evaluate system adoption, business value, and configuration effort for SMEs, low-AI-maturity organizations, and non-technical users. Further development will also focus on stronger validation, confidence scoring, complex multimodal documents, and closer integration with enterprise systems.

8 Acknowledgements

The research was supported by ERDF (European Regional Development Fund).

9 Ethics Declaration

This study did not require formal ethical clearance.

10 AI Declaration

AI tools (Grammarly and ChatGPT) were used during the preparation of this paper to improve grammar, clarity, and written language. The intellectual content and final decisions on the text remained the responsibility of the authors. All AI-generated suggestions were reviewed, edited, and approved by the authors before inclusion in the submitted paper.

References

Alavi, M., Leidner, D.E., 2001. Review: Knowledge Management and Knowledge Management Systems: Conceptual Foundations and Research Issues. MIS Quarterly 25, 107–136. https://doi.org/10.2307/3250961

Amazon Web Services, 2026. Amazon Textract Features. https://aws.amazon.com/textract/features/

Anthropic, 2026. Structured Outputs. https://platform.claude.com/docs/en/build-with-claude/structured-outputs

Ateia, S., Kruschwitz, U., Scholz, M., Koschmider, A., Almohaishi, M., 2025a. LLM-Based Information Extraction to Support Scientific Literature Research and Publication Workflows, in: New Trends in Theory and Practice of Digital Libraries, Communications in Computer and Information Science. Springer, Cham. https://doi.org/10.1007/978-3-032-06136-2_9

Ateia, S., Kruschwitz, U., Scholz, M., Koschmider, A., Almohaishi, M., 2025b. LLM-Based Information Extraction to Support Scientific Literature Research and Publication Workflows, in: New Trends in Theory and Practice of Digital Libraries, Communications in Computer and Information Science. Springer, Cham. https://doi.org/10.1007/978-3-032-06136-2_9

Dagdelen, J., Dunn, A., Lee, S., Walker, N., Rosen, A.S., Ceder, G., Persson, K.A., Jain, A., 2024a. Structured Information Extraction from Scientific Text with Large Language Models. Nat. Commun. 15, 1418. https://doi.org/10.1038/s41467-024-45563-x

Dagdelen, J., Dunn, A., Lee, S., Walker, N., Rosen, A.S., Ceder, G., Persson, K.A., Jain, A., 2024b. Structured Information Extraction from Scientific Text with Large Language Models. Nat. Commun. 15, 1418. https://doi.org/10.1038/s41467-024-45563-x

Finnish AI Region, 2026. Finnish AI Region. https://www.fairedih.fi/en/frontpage/

GAIK Project, 2026a. GAIK: Generative AI Knowledge Management Toolkit. https://github.com/GAIK-project/gaik-toolkit/tree/main

GAIK Project, 2026b. Extractor. https://github.com/GAIK-project/gaik-toolkit/tree/main/implementation_layer/src/gaik/software_components/extractor

Google Cloud, 2026. Document AI. https://cloud.google.com/document-ai

Jarrahi, M.H., Askay, D., Eshraghi, A., Smith, P., 2023. Artificial Intelligence and Knowledge Management: A Partnership between Human and AI. Bus. Horiz. 66, 87–99. https://doi.org/10.1016/j.bushor.2022.03.002

Kudryavtsev, D., Khan, U., Kauttonen, J., 2024. Transforming Knowledge Management Using Generative AI: From Theory to Practice, in: Proceedings of the 16th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management – Volume 3: KMIS. SciTePress, pp. 362–370. https://doi.org/10.5220/0013071400003838

Kudryavtsev, D., Khan, U.A., Remes, J., Kauttonen, J., 2025. Reuse and Guidance for Generative AI Solution Development and Implementation: Knowledge Management Perspective, in: Proceedings of the EDOC-CBI Forum 2025: 29th International Conference on Enterprise Design, Operations, and Computing & 27th International Conference on Business Informatics. Lisbon, Portugal.

LangChain, 2026. Structured Output. https://docs.langchain.com/oss/python/langchain/structured-output

Microsoft Azure, 2026. Azure Document Intelligence in Foundry Tools. https://azure.microsoft.com/en-us/products/ai-foundry/tools/document-intelligence

OpenAI, 2026. Structured Model Outputs. https://developers.openai.com/api/docs/guides/structured-outputs

Peffers, K., Tuunanen, T., Rothenberger, M.A., Chatterjee, S., 2007. A Design Science Research Methodology for Information Systems Research. Journal of Management Information Systems 24, 45–77. https://doi.org/10.2753/MIS0742-1222240302

Walsh, J.P., Ungson, G.R., 1991. Organizational Memory. Academy of Management Review 16, 57–91. https://doi.org/10.5465/AMR.1991.4278992

Xu, D., Chen, W., Peng, W., Zhang, C., Xu, T., Zhao, X., Wu, X., Zheng, Y., Wang, Y., Chen, E., 2024a. Large Language Models for Generative Information Extraction: A Survey. Front. Comput. Sci. 18, 186357. https://doi.org/10.1007/s11704-024-40555-y

Xu, D., Chen, W., Peng, W., Zhang, C., Xu, T., Zhao, X., Wu, X., Zheng, Y., Wang, Y., Chen, E., 2024b. Large Language Models for Generative Information Extraction: A Survey. Front. Comput. Sci. 18, 186357. https://doi.org/10.1007/s11704-024-40555-y

Zhang, B., Acosta, S.G., Carlson, P., Bron, S., Doulcet, P.-L., Ospina, D.B., Suo, S., 2026. ParseBench: A Document Parsing Benchmark for AI Agents.

Zhang, Q., Huang, V.S.-J., Wang, B., Zhang, J., Wang, Z., Liang, H., Wang, S., Lin, M., He, C., Zhang, W., 2024. Document Parsing Unveiled: Techniques, Challenges, and Prospects for Structured Information Extraction.

Behaviour as Knowledge: Neuro-Immersive AI Assessment in HRM

Laura Iacovone
University of Milan, Milan, Italy

Abstract

In contemporary organizations, Human Resource Management (HRM) increasingly relies on assessment processes to evaluate leadership potential, behavioural competences, and alignment with organizational values. However, traditional assessment tools—such as self-report questionnaires, structured interviews, and role-playing exercises—primarily capture deliberate and consciously controlled responses, leaving a significant portion of behaviour underexplored, particularly under conditions of ambiguity and time pressure.

This paper proposes a shift in perspective by conceptualizing assessment as a knowledge creation process, in which behaviour is interpreted as a form of tacit knowledge. Drawing on Knowledge Management (KM) theory, the study argues that many critical competences—such as ethical judgment, decision-making under pressure, and interpersonal sensitivity—emerge through action rather than articulation and therefore cannot be fully captured through conventional HR instruments. To address this limitation, the paper explores the role of neuro-immersive technologies, integrating Virtual Reality (VR) and Artificial Intelligence (AI), as assessment infrastructures capable of eliciting and capturing behavioural knowledge in context. Through immersive simulations enriched with behavioural triggers and AI-driven analytics, these systems allow organizations to observe spontaneous decision-making patterns and transform them into actionable knowledge.

Empirically, the study adopts a qualitative multiple-case approach, analysing two neuro-immersive assessment implementations conducted in large Italian organizations, involving approximately 200 participants. The first case focuses on value-based assessment, highlighting discrepancies between declared organizational values and actual behaviours. The second case examines behavioural potential assessment, revealing latent dimensions of individual performance and development needs.

The findings support a conceptual model in which simulation design variables activate cognitive-emotional processes, leading to observable behavioural outputs that can be externalized as organizational knowledge. The study contributes to KM and HRM literature by reframing assessment as a mechanism for knowledge generation and organizational alignment.

From a practical perspective, the paper provides insights into how immersive technologies can enhance HR practices while emphasizing the importance of ethical governance, transparency, and reflective feedback processes. Ultimately, the study argues that sustainable organizations depend not only on technological advancement but on their ability to transform behavioural evidence into shared knowledge.

Keywords: Organizational behaviour, Knowledge management, Tacit knowledge, Neuro-immersive technologies, HRM assessment

Paper type: Academic Research Paper

1 Introduction

In recent years, immersive technologies such as Virtual Reality (VR) and Artificial Intelligence (AI) have gained increasing relevance within organizational contexts. To date, most existing applications have focused on training and behavioural change, particularly in areas such as leadership development and DE&I, where they are mainly used to enhance experiential learning.

However, despite their growing adoption, their potential as *assessment infrastructures* remains underexplored. Recent studies on AI-driven HRM and algorithmic management further highlight the growing importance of behavioural data in organizational decision-making, while raising critical concerns related to transparency, bias, and governance (Bandara et al., 2025; Milanez et al., 2025).

This paper introduces a shift in perspective: rather than considering immersive technologies as tools for training and sensitization, it conceptualizes them as *mechanisms for capturing behavioural knowledge in action.*

This shift is particularly relevant in the context of Human Resource Management (HRM), where traditional assessment methods—such as structured interviews, self-report questionnaires, psychometric tests, and role-playing exercises—are inherently limited in their ability to capture spontaneous, context-dependent, and unconscious behavioural dynamics. These methods primarily rely on declared intentions and controlled responses, while many critical competences—such as ethical judgment, people orientation, and decision-making under pressure—emerge in real-time under conditions of pressure, ambiguity, and social interaction, often outside conscious awareness.

From a Knowledge Management (KM) perspective, this limitation reflects a deeper epistemological issue: organizations lack effective mechanisms to capture, externalize, and reuse *tacit behavioural knowledge* (Nonaka

& Takeuchi, 1995). Behaviour, particularly under conditions of stress, ambiguity, and social pressure, can be interpreted as situated knowledge that remains largely invisible within existing assessment systems.

This paper addresses this gap by proposing a *knowledge-based reconceptualization of assessment*, in which immersive neuro-technological environments enable the activation, observation, and interpretation of behavioural patterns. Through AI-driven analytics and structured feedback processes, these behavioural traces can be transformed into actionable organizational knowledge assets.

Empirically, the study is based on two neuro-immersive assessment implementations conducted in large Italian organizations, involving approximately 200 participants. These cases provide a unique opportunity to explore how immersive simulations can reveal discrepancies between declared values and enacted behaviours, as well as to identify latent dimensions of individual potential. The study proposes a conceptual model in which immersive simulation design elements activate behavioural responses through cognitive and emotional mediators, enabling the externalization of tacit knowledge in the form of observable decision patterns.

The paper contributes to the IFKAD 2026 theme by demonstrating how intelligent technologies can support the transformation of behavioural evidence into knowledge, thus enabling more sustainable and reflective organizational systems.

2 Theoretical Background

2.1 Behaviour as tacit knowledge

Knowledge Management literature has long distinguished between *explicit* and *tacit* knowledge, emphasising that a significant portion of what individuals know cannot be easily articulated or codified (Nonaka & Takeuchi, 1995). Tacit knowledge is deeply embedded in action, experience, and context, and is often revealed through practice rather than verbalisation.

Within organizational settings, many critical competences—such as leadership, decision-making, ethical judgement, and interpersonal sensitivity—are inherently tacit. These competences are not only difficult to formalise, but also highly dependent on situational dynamics, making them particularly challenging to assess through traditional HRM tools.

The process of knowledge creation, as described in the SECI model (Nonaka & Takeuchi, 1995), and further developed in studies on knowledge processes and organizational use of knowledge (Davenport & Prusak, 1998), highlights the importance of externalising tacit knowledge through interaction, reflection, and shared experience. However, while considerable attention has been devoted to how knowledge is created and shared, less emphasis has been placed on how tacit knowledge can be *captured and interpreted in real-time behavioural contexts.*

From this perspective, *behaviour* can be understood as a manifestation of *tacit knowledge in action*. Rather than being treated solely as an outcome to be evaluated, behaviour represents an observable expression of how individuals interpret situations, prioritise actions, and respond to environmental stimuli.

Organizational learning theory further supports this view by emphasising the role of action and reflection in the development of knowledge (Argyris & Schön, 1978; Davenport & Prusak, 1998). In particular, the distinction between espoused theories and theories-in-use suggests that what individuals claim to believe may differ from how they actually behave in practice. This gap is especially relevant in assessment contexts, where declared values and observed behaviours may not align.

Despite these theoretical insights, traditional assessment methods remain largely focused on explicit knowledge and self-reported information, limiting their ability to capture the tacit dimension of competences. As a result, organizations may base critical decisions on incomplete or biased representations of individual capabilities.

This limitation of traditional HRM assessment methods, which rely on verbalization and abstraction, calls for new approaches capable of accessing tacit knowledge through observation of behaviour in context. In this sense, immersive environments offer a promising avenue, as they enable the creation of realistic situations in which tacit knowledge can be enacted, observed, and subsequently interpreted.

This perspective provides the foundation for interpreting behaviour not merely as an outcome of action, but as a source of knowledge that can be systematically observed and analysed.

2.2 Decision-making and automatic behaviour

Understanding human behaviour in organizational contexts requires moving beyond purely rational models of decision-making. A substantial body of research in cognitive psychology and neuroscience has demonstrated

that individuals often rely on automatic, intuitive processes, particularly under conditions of uncertainty, time pressure, and emotional activation (Kahneman, 2011; Damasio, 2018).

According to dual-process theories, decision-making operates through two interacting systems: a fast, automatic, and intuitive system (System 1), and a slower, deliberate, and analytical system (System 2) (Kahneman, 2011). While traditional assessment methods tend to capture responses associated with deliberate reasoning, many critical behaviours in organizational settings are driven by automatic processes that emerge in real-time interactions.

These automatic responses are particularly relevant in contexts characterized by ambiguity, competing priorities, and social complexity, where individuals are required to make rapid decisions without the possibility of extended reflection. In such situations, behaviour reflects not only cognitive evaluation but also emotional states, prior experiences, and implicit assumptions, making it a rich source of tacit knowledge.

From an organizational perspective, this has important implications for assessment practices. If decision-making is significantly influenced by automatic processes, then methods based solely on self-report or structured evaluation may provide only a partial representation of individual competences. This limitation is especially critical when assessing soft skills such as leadership, responsibility, and interpersonal sensitivity, which are often enacted rather than explicitly articulated.

Immersive environments offer a unique opportunity to address this gap. By recreating realistic scenarios and introducing elements of unpredictability and pressure, they activate cognitive-emotional processes that are more likely to trigger spontaneous behavioural responses. In this sense, neuro-immersive simulations enable the observation of decision-making as it unfolds, providing access to behavioural dynamics that remain largely inaccessible in traditional assessment settings.

2.3 Immersive technologies and behavioural observation

While immersive technologies have been widely explored in training contexts, their application as assessment infrastructures remains limited, particularly in relation to behavioural data interpretation and knowledge generation.

Recent literature confirms that immersive technologies, particularly Virtual Reality, are increasingly used for soft-skill development and experiential learning, although primarily within training-oriented frameworks rather than assessment contexts (Dubiel et al., 2025).

At the same time, emerging research highlights the potential of immersive environments as platforms for behavioural measurement, enabling the capture of stress responses and decision-making dynamics in real time (Rah & Chen, 2025).

This evolution is particularly relevant in light of broader developments in AI-driven HRM and algorithmic management, which emphasise the growing role of behavioural data while also raising concerns related to bias, transparency, and ethical governance (Bandara et al., 2025; Milanez et al., 2025; Naoum et al., 2026).

Immersive environments enable the simulation of realistic contexts in which individuals interact naturally. When combined with AI-based analytics, these environments allow for the systematic observation of behavioural patterns, opening new possibilities for assessment. This reinforces the need to investigate how immersive environments may function not only as learning tools, but also as infrastructures for *behavioural observation* and *knowledge generation*.

3 Conceptual Framework

Building on the integration between Knowledge Management, organizational behaviour, and immersive technologies, this study proposes a conceptual model that explains how neuro-immersive environments enable the externalization of tacit behavioural knowledge and the transformation of behavioural evidence into organisational knowledge. Based on the proposed framework, the study advances the following propositions:

- **P1**: Simulation design variables positively influence cognitive-emotional activation.
- **P2**: Cognitive-emotional activation facilitates the emergence of spontaneous behavioural responses.
- **P3**: Behavioural responses can be captured and interpreted as forms of tacit knowledge.

The model is structured as a dynamic flow system consisting of four main dimensions, as illustrated in Figure 1:

1. **Simulation Design Variables (Input Layer)**
    - Scenario realism
    - Behavioural triggers (e.g., unexpected events, ambiguity, social tension)
    - AI-driven interaction (adaptive responses, conversational agents)

2. **Cognitive-Emotional Mediators (Process Layer)**
    - Emotional activation
    - Cognitive dissonance
    - Decision-making under pressure

These mediators play a crucial role in reducing conscious control and activating automatic behavioural responses (Kahneman, 2011; Damasio, 2018).

3. **Behavioural Knowledge Outputs (Output Layer)**
    - Observable behavioural patterns
    - Decision-making pathways
    - Tacit knowledge externalization
4. **Organizational Knowledge (Outcome Layer)**
    - Feedback
    - HR insights
    - KM integration

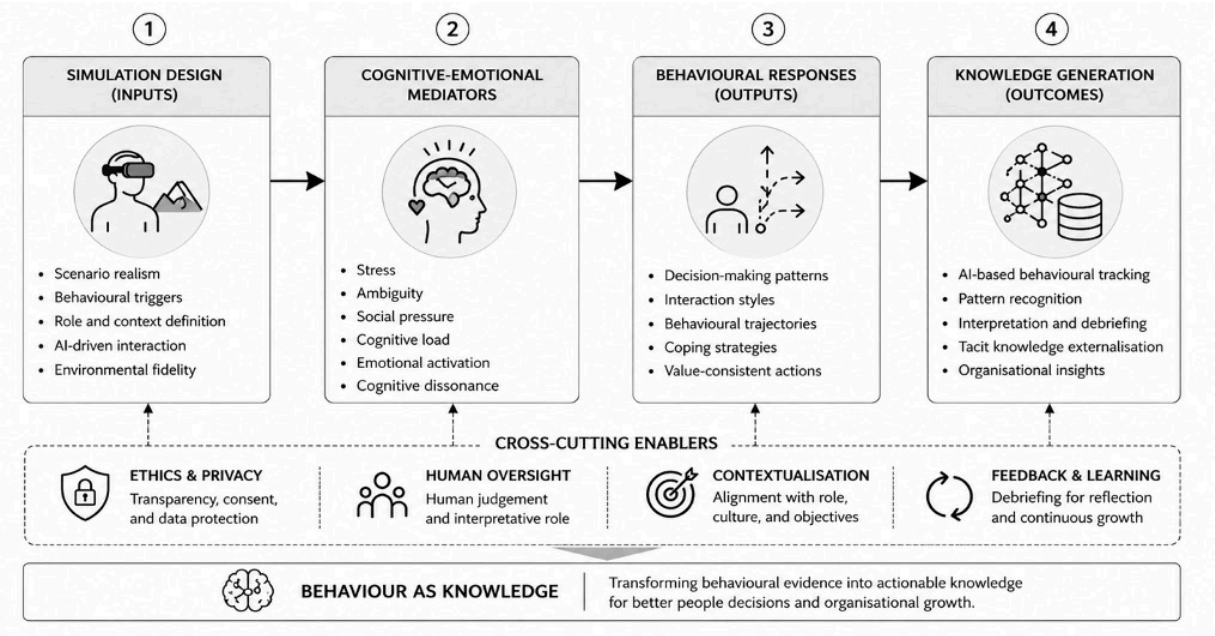


Figure 1. Neuro-Immersive Behavioural Knowledge Generation Model

The proposed model conceptualises neuro-immersive assessment as a process through which behavioural knowledge is generated and formalised. It is structured as a multi-layered system linking simulation design, cognitive-emotional activation, behavioural responses, and knowledge outcomes. Through AI-based tracking systems, these outputs are captured, structured, and interpreted, enabling their transformation into organizational knowledge.

At the *input level*, simulation design elements—such as scenario realism, behavioural triggers, and AI-driven interaction—create a controlled yet ecologically valid environment.

These elements activate a set of *cognitive-emotional mediators*, including stress, ambiguity, and social pressure, which are known to influence decision-making processes (Kahneman, 2011; Damasio, 2018).

These mediators reduce the extent of deliberate control and facilitate the emergence of spontaneous *behavioural responses*. Such responses manifest as observable decision-making patterns, interaction styles, and behavioural trajectories.

Finally, through AI-based tracking and structured feedback mechanisms, these behavioural outputs are captured, interpreted, and transformed into *organisational knowledge*, supporting learning and alignment. In this sense, the model positions assessment not as a static evaluation process, but as a dynamic system of knowledge generation and interpretation.

4 Methodology

This study adopts a qualitative multiple-case research design aimed at analytical generalisation (Eisenhardt, 1989; Yin, 2018), combined with an abductive research approach that iteratively connects empirical observations with theoretical constructs.

The empirical setting includes two neuro-immersive assessment implementations conducted in large Italian organisations operating in the consulting and financial services sectors, involving a total of approximately 200 participants (130 partners and senior managers, and 70 high-potential talents).

The selection of these cases follows a theoretical replication logic, as they represent distinct organisational contexts and assessment objectives—namely value alignment and behavioural potential evaluation—thereby enabling cross-case comparison and enhancing the robustness of analytical generalisation.

Data collection relied on a multi-source triangulation strategy, including direct observation of immersive sessions, behavioural data generated by AI-based decision-tracking systems, scenario design documentation and underlying algorithmic logic, individual feedback reports and aggregated analytics, debriefing sessions, and facilitated group discussions.

Data analysis was conducted through a three-stage qualitative coding process (open, axial, and selective coding), supported by behavioural analytics: a) *Open coding*, identifying recurring behavioural patterns and decision-making dynamics across scenarios; b) *Axial coding,* linking observed behaviours to contextual triggers (e.g., stress, ambiguity, social pressure); c) *Selective coding*, mapping behavioural patterns to higher-level constructs related to tacit knowledge and value alignment.

In parallel, the analysis also integrates *behavioural analytics outputs*, including decision sequences, interaction patterns, and response variability, to support interpretation and reduce reliance on self-reported data.

This approach enabled the identification of recurring behavioural patterns and their interpretation as forms of tacit knowledge, in relation to contextual triggers and higher-level constructs. To enhance robustness, cross-case analysis was performed to identify convergent and divergent behavioural dynamics across the two contexts.

Although participants were aware of being observed, the immersive and time-constrained nature of the simulations reduced the degree of deliberate control, allowing more spontaneous behavioural responses to emerge. Scenario standardisation ensured comparability across participants, while embedded unexpected events introduced sufficient variability to capture individual differences.

This methodological approach enables the systematic transformation of behavioural evidence into structured organisational knowledge, supporting the study's conceptualisation of assessment as a knowledge creation process.

5 Results

5.1 Case 1: Value-based assessment and behavioural alignment

The first case of neuro-immersive simulations focuses on the assessment of alignment between declared organizational values and actual managerial behaviour. Participants—all senior managers—were immersed in simulations reproducing a full working day, including both professional and relational situations. These scenarios encompassed interactions with colleagues, decision-making under time constraints, conflict management, and situations requiring prioritization between organizational goals and interpersonal considerations.

The design of the simulations intentionally introduced elements of ambiguity, competing priorities, and social tension, allowing participants to navigate situations in which no single "correct" answer was available. This structure made it possible to observe not only explicit decision outcomes, but also *how attention was distributed between people*, *tasks*, *and objectives*, both at a conscious and implicit level.

From a behavioural perspective, the presence of unexpected events and social tension acted as *triggers activating automatic responses*, often bypassing conscious control. These responses were captured through AI-based behavioural tracking, revealing recurring decision-making patterns across participants.

The analysis revealed a consistent gap between *declared values* (e.g., inclusivity, responsibility, collaboration) and *observed behavioural patterns* under pressure. While participants were generally able to articulate organizational values, their behaviour in immersive contexts often reflected alternative priorities, such as efficiency over listening, risk aversion over accountability, and goal achievement over relational sensitivity.

The results showed a high degree of variability across participants, leading to the identification of *distinct behavioural clusters*, characterised by different orientations toward people, performance, and responsibility. These clusters did not emerge from predefined categories, but from recurring behavioural patterns observed across scenarios.

Importantly, the structured debriefing process enabled participants to reflect on these discrepancies and patterns, transforming individual behavioural evidence into *shared insights* and *collective knowledge about organizational culture.*

5.2 Case 2: Behavioural potential assessment

The second case examines the assessment of behavioural potential within a financial institution, focusing on a group of high-potential talents identified for future leadership roles. Participants engaged in immersive simulations structured as a dynamic working-day narrative, including multiple decision points, time constraints, unexpected events, and evolving interpersonal interactions.

The simulations were designed to reflect both operational and relational dimensions of work, requiring participants to balance task execution, decision-making under pressure, and attention to colleagues and stakeholders. This holistic structure enabled the observation of behaviour across a range of contexts, rather than in isolated or artificial tasks.

The findings show that neuro-immersive environments are particularly effective in revealing *latent dimensions of behavioural potential*, which are not easily captured through traditional assessment methods.

The analysis highlighted the emergence of *distinct behavioural trajectories*, reflecting different ways of interpreting complexity and responding to uncertainty. AI-based analytics enabled the identification of recurring patterns, as well as the detection of inconsistencies between self-perception and observed behaviour.

Participants showed significant variation in their ability to manage pressure, prioritise effectively and responsibly, and maintain relational awareness (listening and interpersonal sensitivity) while pursuing performance objectives.

Notably, the analysis revealed that behavioural potential could not be reduced to a single dimension, but rather emerged as a combination of patterns, including adaptability to unexpected situations, consistency, and responsiveness to context. In several cases, discrepancies were observed between participants' self-perception and their actual behaviour in the simulations, providing valuable insights for development.

Similarly to the first case, the identification of recurring behavioural patterns enabled the classification of participants into *behavioural clusters*, reflecting different orientations toward leadership, decision-making under pressure, and interpersonal engagement.

The combination of immersive experience, behavioural feedback, and identified clusters allowed for a more nuanced understanding of individual potential, supporting HR processes such as talent development and succession planning.

5.3 Cross-case insights: from behavioural data to organizational knowledge

An important contribution of this study concerns the transformation of behavioural data into organizational knowledge. The findings suggest that the value of neuro-immersive assessment does not lie solely in the collection of behavioural data, but in the processes through which these data are interpreted, structured, and restituted within the organization.

Across both cases, immersive simulations generated rich behavioural evidence in the form of decision patterns, interaction styles, and response trajectories. Specifically, simulation design variables (such as realism, triggers, and AI interaction) consistently activated behavioural responses, while cognitive-emotional mediators—including stress, ambiguity, and social pressure—reduced conscious control and facilitated spontaneous actions. These dynamics resulted in behavioural outputs observable as recurring decision-making patterns.

Due to confidentiality constraints, company names and raw data are not disclosed; however, the consistency of behavioural patterns across participants supports the robustness of the findings.

These behavioural traces, when analysed systematically, revealed consistent differences among participants, enabling the identification of *distinct behavioural clusters*. Such clusters reflect different orientations toward key organizational dimensions, including the balance between task achievement and attention to people, as well as the ability to manage complexity and uncertainty.

The identification of these clusters represents a crucial step in transforming raw behavioural data into meaningful knowledge. Rather than focusing on isolated actions, the analysis highlights recurring patterns that can be interpreted as expressions of tacit knowledge in action. In this sense, behavioural data become *knowledge artefacts*, capturing how individuals actually interpret situations, prioritise objectives, and respond to contextual demands, enabling organizations to identify behavioural tendencies, detect misalignments between values and actions, and generate actionable insights for HRM.

This transformation process is further reinforced through structured feedback and debriefing activities, in which individual behavioural patterns are discussed, compared, and collectively interpreted. Through this process, personal experiences are translated into shared reflections, enabling the emergence of insights related to organizational culture, leadership styles, and development needs.

Importantly, this approach differentiates neuro-immersive assessment from forms of algorithmic management focused on monitoring and control. Rather than automating decisions, the system supports *human-mediated interpretation*, ensuring that behavioural data are used as a basis for reflection and learning. This is consistent with recent research highlighting the importance of transparency, ethical governance, and the responsible use of AI in HRM contexts (Bandara et al., 2025; Milanez et al., 2025).

Overall, the findings demonstrate that immersive assessment environments can act as *knowledge-generating systems*, enabling organizations to capture, interpret, and reuse behavioural evidence in a structured and meaningful way.

6 Discussion

They extend the literature by extending the conceptualisation of assessment from a *measurement-oriented activity* to a *knowledge-generative process* grounded in *behavioural evidence.*

6.1 Assessment as knowledge creation

From a Knowledge Management perspective, this shift is particularly relevant. The findings contribute to a reconceptualization of assessment within HRM. The results support the view that behaviour, especially when enacted under conditions of ambiguity and pressure, represents a form of tacit knowledge that cannot be fully captured through traditional assessment methods. By enabling the observation of behaviour in context, neuro-immersive environments provide access to dimensions of knowledge that would otherwise remain implicit and inaccessible. Rather than being limited to performance measurement, assessment can be understood as a *knowledge creation process*, in which *behavioural evidence is generated*, *interpreted*, *and reintegrated into organizational systems.*

This perspective aligns with KM theory, particularly the notion that tacit knowledge can be externalized through experience and reflection (Nonaka & Takeuchi, 1995). Neuro-immersive environments extend this process by providing structured contexts in which behaviour can be observed under realistic conditions, thus bridging the gap between knowing and doing.

This interpretation extends traditional Knowledge Management perspectives by integrating behavioural evidence into knowledge creation dynamics, reinforcing the role of action-based learning processes (Davenport & Prusak, 1998).

6.2 The Role of cognitive-emotional mediators

A central insight emerging from the analysis concerns the role of *cognitive-emotional mediators*. The simulations show that factors such as time pressure, social interaction, and uncertainty significantly influence how individuals act, often reducing the extent of deliberate control and activating automatic decision-making processes (Kahneman, 2011; Damasio, 2018). These mediators are critical, as they reduce the reliability of self-reported data, reveal behavioural patterns that remain hidden in controlled settings, and enable a more authentic assessment of soft skills.

This highlights the importance of designing assessment environments that are not only realistic, but also *intentionally disruptive*, in order to activate meaningful behavioural responses. This reinforces the idea that assessment environments should not aim to eliminate complexity, but rather to incorporate it as a necessary condition for revealing meaningful behavioural patterns.

The empirical evidence also highlights the importance of moving beyond isolated observations toward the identification of recurring behavioural configurations. The emergence of distinct behavioural clusters across participants suggests that individuals tend to adopt relatively stable orientations in the way they balance competing demands, such as task achievement and attention to people. These orientations, while not always consciously articulated, represent valuable insights into how individuals interpret and navigate organizational contexts.

6.3 From behavioural data to organizational knowledge

Another important contribution concerns the transformation of *behavioural data* into *organizational knowledge*. The study shows that the value of immersive assessment lies not only in data collection, but in the processes of interpretation and restitution that follow.

Through structured debriefing and collective reflection, individual behavioural patterns are translated into shared reflections, cultural insights and development pathways. This process transforms behavioural outputs into *knowledge assets*, contributing to organizational learning and alignment. In this sense, assessment becomes a mechanism through which organizations can make sense of their own behavioural dynamics.

At the same time, the increasing use of AI-driven systems in HRM raises important ethical considerations. As highlighted by recent research on algorithmic management, the use of behavioural data may lead to risks related to algorithmic bias, opacity, and unintended consequences of data-driven decision-making processes or forms of control (Floridi, 2021; Bandara et al., 2025; Milanez et al., 2025).

Importantly, the neuro-immersive approach proposed in this study emphasizes *reflection*, *transparency*, *and developmental use of data*. Behavioural evidence is *not used to automate decisions*, *but to support dialogue and learning*.

6.4 Implications for HRM and organizational development

From a managerial perspective, the findings suggest that neuro-immersive assessment can complement traditional HR tools by enabling organizations to enhance talent identification, support leadership development, identify gaps between organizational values and actual practices, and foster a more coherent organizational culture.

At the same time, the adoption of such systems requires careful attention to governance issues, including transparency of evaluation criteria, ethical use of behavioural data, and avoidance of bias in AI-based analytics.

Overall, the findings suggest that neuro-immersive assessment can play a key role in bridging the gap between *behaviour* and *knowledge*. By enabling the observation of how individuals act in complex situations, these environments support a more nuanced and dynamic understanding of competences, contributing to more informed and sustainable organizational practices.

7 Conclusion

This study has explored the potential of neuro-immersive technologies as assessment infrastructures capable of capturing and transforming behavioural knowledge within organizational contexts. By integrating insights from Knowledge Management, organizational behaviour, and emerging technologies, the paper has proposed a conceptual shift *from assessment as measurement* to *assessment as knowledge creation*.

The findings demonstrate that immersive environments enable the observation of behaviour under realistic conditions, revealing patterns that are often inaccessible through traditional assessment methods. In doing so, they provide a means of *externalising tacit knowledge*, contributing to a more comprehensive understanding of individual competences and organizational dynamics.

Importantly, the study highlights that the value of such approaches does not lie solely in technological innovation, but in the ability to interpret behavioural evidence and integrate it into organizational learning processes. In this sense, neuro-immersive assessment should not be understood as a replacement for existing tools, but as a complementary layer that enriches the understanding of behaviour in context.

From a practical perspective, the adoption of immersive assessment systems offers new opportunities for HRM, particularly in areas such as talent development, leadership evaluation, and cultural alignment. However, these opportunities must be balanced with careful attention to ethical considerations, including transparency, fairness, and the responsible use of behavioural data.

The study is not without limitations. The qualitative nature of the research and the limited number of cases constrain the generalizability of the findings. Furthermore, the interpretation of behavioural data remains dependent on contextual and organizational factors, requiring careful and informed analysis.

Future research could extend this work by exploring the application of the proposed model across different sectors (e.g., pharmacy counselling simulations), integrating quantitative behavioural metrics, and examining the long-term impact of immersive assessment on organizational culture and performance.

Ultimately, this paper argues that the sustainability of intelligent organizations depends not only on their ability to generate data, but on their capacity to transform behavioural evidence into meaningful knowledge. In this perspective, assessment becomes not merely a tool for evaluation, but a strategic process for understanding and shaping organizational behaviour.

The results therefore suggest that neuro-immersive assessment environments enable the *externalization of tacit knowledge*, supporting its integration into organizational learning processes.

8 Ethics Declaration

This study was conducted within private organizational contexts and did not involve the collection of sensitive personal data. All data were fully anonymized, and company names have been omitted to ensure confidentiality.

Given the nature of the study and the absence of sensitive data, vulnerable populations or experimental interventions, formal ethical approval was not required. However, the research was conducted in accordance with general ethical principles of academic research, including respect for participants, data confidentiality, and responsible data use.

9 AI Declaration

Generative AI (ChatGPT, OpenAI, April 2026 version) was used to support preliminary drafting in translation refinement, and the retrieval of non-academic sources (e.g., corporate case studies and reports). All content was subsequently verified, reviewed, and critically validated by the author.

References

Argyris, C., & Schön, D. A. (1978) Organizational learning: A theory of action perspective. Addison-Wesley.

Bandara, R. J., Biswas, K., Akter, S., Shafique, S., & Rahman, M. (2025) Addressing algorithmic bias in AI-driven HRM systems: Implications for strategic HRM effectiveness, *Human Resource Management Journal*, *35*(4), 1047–1063. https://doi.org/10.1111/1748-8583.12609.

Damasio, A. (2018) *The strange order of things: Life, feeling, and the making of cultures*, Pantheon Books.

Davenport, T. H., & Prusak, L. (1998) Working knowledge: How organizations manage what they know. Harvard Business School Press.

Dubiel, A., Kamińska, D., Zwoliński, G., Ramić-Brkić, B., Agostini, D., & Zancanaro, M. (2025) Virtual reality for the training of soft skills for professional education: Trends and opportunities, *Interactive Learning Environments*, 1–21. https://doi.org/10.1080/10494820.2025.2450634

Eisenhardt, K. M. (1989) Building theories from case study research, *Academy of Management Review*, 14(4), 532–550. Floridi, L. (2021). *Ethics of artificial intelligence.* Oxford University Press.

Kahneman, D., (2011) *Thinking, Fast and Slow*, Farrar, Straus and Giroux

Milanez, A., Lemmens, A., & Ruggiu, C. (2025) Algorithmic management in the workplace. OECD Publishing. https://doi.org/10.1787/287c13c4-en

Naoum, R. F., Szakadáti, T., & Balogh, G. (2026) Artificial Intelligence (AI) in human resource management (HRM): A systematic review of its dual impact on diversity, equity, and inclusion (DEI), *Management Review Quarterly*. https://doi.org/10.1007/s11301-025-00580-y.

Nonaka, I., Takeuchi, H. (1995) *The Knowledge-Creating Company: how Japanese Companies create the Dynamics of Innovation*, Oxford University Press

Rah, A., & Chen, Y. (2025) Virtual reality as a stress measurement platform: Real-time behavioral analysis with minimal hardware, *Sensors*, *25*(17), 5323. https://doi.org/10.3390/s25175323.

Úbeda-García, M., Marco-Lajara, B., Zaragoza-Sáez, P. C., & Poveda-Pareja, E. (2025) Artificial intelligence, knowledge and human resource management: A bibliometric analysis (2002–2024), *Journal of Innovation & Knowledge*. https://doi.org/10.1016/j.jik.2025.100809

Yin, R. K. (2018) Case study research and applications, Sage.

Individual Intellectual Capital for Sustainability

Jonna Käpylä
Tampere University, Tampere, Finland

Abstract

Given the urgent need for radical sustainability transformation, all fields of research, including intellectual capital theory, must critically examine their underlying foci, points of departure, and assumptions. Previous research on sustainability and intellectual capital has primarily conceptualised sustainable or green intellectual capital at the organisational level, while largely neglecting the individual level and failing to specify the content of individual intellectual capital for sustainability. In response, this study addresses the research question of how individual intellectual capital for sustainability can be defined and conceptualised. The study develops a conceptualisation of individual intellectual capital for sustainability through an integrative review method, in which existing literature is synthesised to generate new theoretical knowledge. Drawing on individual intellectual capital theory as an organising framework, the review integrates insights from research on education for sustainable development, sustainability transformation, sustainability science, and sustainability leadership. The study makes two main contributions. First, it defines individual intellectual capital for sustainability as an asset encompassing the full range of individual cognitive, socio-emotional, and behavioural knowledge and knowing capabilities that contribute to the development of sustainability-oriented thinking, identity, and agency, and thereby promote sustainability transformation. Second, it delineates the core components of individual intellectual capital for sustainability. Individual intellectual capital for sustainability is conceptualised as a multidimensional asset comprising human, structural, and relational capital, each expressed through cognitive (thinking), socio-emotional (identity), and behavioural (agency) domains. Together, these dimensions capture how sustainability-related knowledge, values, relationships, and action capabilities are embodied in individuals in ways that enable sustainability transformation. For example, systems thinking, an ecologically grounded worldview, and a personal connection to sustainability are identified as key intangible assets for advancing sustainability. The resulting integrated framework provides a foundation for the operationalisation and measurement of individual intellectual capital for sustainability and offers guidance for human resource management aimed at advancing sustainability transformation.

Keywords: Individual intellectual capital, Sustainable development, Individual intellectual capital for sustainability, Sustainability competencies, Integrative review.

Paper type: Academic Research Paper

1 Introduction

Humanity is currently operating outside the safe planetary boundaries; six of the nine planetary boundaries, including biosphere integrity, climate change and novel entities, have already been crossed (Richardson et al., 2023; Stockholm Resilience Centre, 2026). This points to the need for radical sustainability transformation in societies, where existing worldviews and paradigms must be questioned and critically evaluated (Abson et al., 2017). Each field of research, including the theory of intellectual capital, must critically examine its own starting points and assumptions. Thus, the only yardstick of intellectual capital management cannot be profit and competitive advantage but also ecological and social value. Čabrilo, Užienė and Kianto (2025) call for better recognition of sustainability in intellectual capital research and aligning it with long-term value creation and societal impact rather than short-term performance. Intellectual capital, as representing the knowledge and knowing capability (Nahapiet and Ghoshal, 1998), has huge transformational potential in contributing to sustainable development. However, this requires that our knowledge and knowing capability serves sustainability and not the unsustainable development trajectories of previous decades, which have led to overconsumption and environmental degradation.

The traditional framework of intellectual capital understands sustainability as one of several intangible assets (Benevene et al., 2021) and not as an overarching theme that is integrated into the concept of intellectual capital. The discussion of the interrelationship between intellectual capital and sustainability has gained ground in recent years (Alvino et al., 2021) and the previous literature on sustainability and intellectual capital has defined and explicated sustainable or green intellectual capital at the organizational level (Chen, 2008) but not at an individual level and furthermore, not specifying the content of individual intellectual capital for sustainability. Therefore, this study addresses the research question of *how to define and conceptualise individual intellectual capital for sustainability.*

The study defines and conceptualises individual intellectual capital for sustainability by applying the method of integrative review where existing literature is used to create new knowledge (Torraco, 2005; 2016). The theory of intellectual capital (Martín-de-Castro et al., 2011) is used as a guiding theory to conceptually structure the topic and organise the synthesis. The critical analysis and synthesis of the key intangible assets for sustainability are conducted by enriching intellectual capital theory with new insights, concepts, and theoretical perspectives drawn from the fields of education for sustainable development, including transformative learning for sustainability (e.g. Rodríguez Aboytes and Barth, 2020) and sustainability competencies (e.g. Redman and Wiek, 2021); sustainability transformation and sustainability science (e.g. Amel et al., 2017); and sustainability leadership (e.g. Sajjad, Eweje and Raziq, 2024). Accordingly, this research aims to further develop the concept of individual intellectual capital by integrating perspectives from related disciplines in which the skills and competencies required for sustainability transformation have been extensively examined. For emerging research topics such as this, the integration of literature across multiple fields can offer novel and valuable perspectives (Torraco, 2016). Although individual-level intellectual capital has gained increasing scholarly attention in recent years (Wang et al., 2020; Khan et al., 2024; Kmieciak, 2024; Alppivuori and Kianto, 2025), research in this area remains limited.

2 Intellectual Capital Theory

2.1 Intellectual capital and sustainable intellectual capital

The concept of intellectual has developed in an organisational context, and many of the definitions of intellectual capital refer to the organisational context (Martín-de-Castro et al., 2011). Meso- and macro levels (i.e. organisations, networks, regions, nations) have dominated the literature on intellectual capital, rather than the micro level of individuals. Intellectual capital has been defined as “the knowledge and knowing capability of a social collectivity” and “a valuable resource and a capability for action based in knowledge and knowing” (Nahapiet and Ghoshal, 1998, p. 245) or “the sum of all knowledge firms utilise for competitive advantage” (Subramaniam and Youndt, 2005) thus emphasising its collective and instrumental nature. In the light of these definitions, intellectual capital can be characterised by three features: intangibility (i.e., knowledge-related assets), the ability to create value now and in the future, and its social or collective nature.

Intellectual capital has been divided into different dimensions the most established being the classification into human, structural and relational capital (Martín-de-Castro et al., 2011). Another well-known and partly similar categorisation is the classification into human, organisational and social capital (Subramaniam and Youndt, 2005; Kang and Snell, 2009). Each aspect accumulates and distributes knowledge differently: human capital through individuals, organisational capital through organisational structures, processes and systems, and social capital through relationships and networks (Subramaniam and Youndt, 2005).

Research on sustainable or green intellectual capital has only recently begun to gain traction. To date, the dominant focus of this body of research has been on examining the effects of green or sustainable intellectual capital on organisational performance (e.g. Yusliza et al., 2020). Alvino et al. (2021) link intellectual capital and sustainability through the notion of long-term value creation, while the review by Benevene et al. (2021) demonstrates that green intellectual capital contributes to enhanced organisational performance and competitive advantage. Although the existence of green intellectual capital at the individual level is acknowledged in the literature, it remains insufficiently elaborated. For example, Chen (2008, p. 277) defines green intellectual capital as “the total stocks of all kinds of intangible assets, knowledge, capabilities, and relationships, etc. about environmental protection or green innovation at the individual level and the organisational level within a company”.

2.2 Individual intellectual capital

While a small number of studies addressed individual-level intellectual capital prior to the 2020s (e.g. Arenas et al., 2013), it is only more recently that this line of research has begun to gain momentum. Despite growing scholarly interest, research on individual intellectual capital remains relatively scarce (Wang et al., 2020; Wang, Bu and Cai, 2021; Bu et al., 2023; Khan et al., 2024; Kmieciak, 2024; Hong, Chen, and Zhang, 2025; Alppivuori and Kianto, 2025). This limited development is well illustrated by search results from the Scopus database. A search for the term “intellectual capital” (in article titles, abstracts, and keywords) yielded approximately 9,200 results (as of 24 April 2026), whereas a search for “individual intellectual capital” returned only 25 results. Given the traditional definitions of intellectual capital, individual intellectual capital is somewhat of an oxymoron,

juxtaposing concepts with opposite meanings: analysing a phenomenon recognised as collective at the individual level.

Individual intellectual capital has been defined by Wang et al. (2020, p. 1107) as "the sum of knowledge and skills owned or controlled by individual employees". Furthermore, Wang, Bu and Cai (2021, p. 1206) emphasise that it "is accumulated through one's intellectual activity, and constitutes competitive advantage for the individual, as well as for his/her organisation". As with conceptualisations at the organisational level, definitions of intellectual capital at the individual level are thus closely linked to the notion of competitive advantage. In a similar vein, Alppivuori and Kianto (2025, p. 67), building on Wang et al. (2020), conceptualise individual intellectual capital as "the sum of knowledge, skills, and abilities held by an individual, which the individual uses to succeed in their work". Taken together, these definitions position individual intellectual capital primarily as an instrumental resource for achieving competitive advantage or success in the workplace.

Individual human capital, individual structural capital and individual relational capital are considered to be the three main sub-dimensions of individual intellectual capital (Wang et al., 2020; Wang, Bu and Cai, 2021). Individual human capital refers to the objective and tactical knowledge accumulated in the process of learning and working (Wang et al., 2020; Wang, Bu and Cai, 2021). It represents work experience, professional skills, academic ability, new ideas and knowledge or innovative consciousness, as operationalised in the measurement tools by Wang et al. (2020) and Wang, Bu and Cai (2021). Individual structural capital refers to the subjective knowledge that one acquires during the thinking and cognitive processes, including the individual's principles, standards, methods and knowledge of business routines (Wang et al., 2020; Wang, Bu and Cai, 2021). In the operationalisation, it represents, for example, ways of thinking, methodologies for dealing with problems, self-management, such as time management and attention to focus on important matters, decision logic and operation procedure (Wang et al., 2020; Wang, Bu and Cai, 2021). Individual relational capital represents an individual's ability to acquire knowledge from various stakeholders, collaborate, interact, communicate, learn from others and build relationships (Wang et al., 2020; Wang, Bu and Cai, 2021). These definitions provide a basis for analysing intellectual capital at the individual level in recent research on the subject (Bu et al. 2023; Khan et al., 2024; Kmieciak, 2024, Hong, Chen and Zhang, 2025). However, several conceptual ambiguities remain. In particular, it remains unclear what is meant by the distinction between objective and subjective knowledge, which is presented as the principal differentiating factor between human and structural intellectual capital. Nor is it evident how the three sub-dimensions differ at a more fundamental level in terms of their underlying knowledge dimensions.

Alppivuori and Kianto (2025) further develop the concept of individual intellectual capital by providing a more detailed structural framework and by specifying its components and sub-elements. They introduce two additional dimensions, emotional capital and renewal capital, and present a five-part model of individual intellectual capital comprising human, relational, structural, renewal, and emotional capital. They do not provide precise definitions of the individual dimensions of; instead, they describe them through their constituent components. For example, emotional capital consists of emotions, recognising emotions, managing emotions, and utilising emotions in action. Furthermore, the five dimensions of individual intellectual capital are classified along two analytical axes: rigid versus dynamic and inward versus outward orientation. In this 2 x 2 matrix, structural capital occupies an intermediate position, functioning as a platform for individual action that enables and supports the effective utilisation of the other components of individual intellectual capital.

Researchers therefore appear to be relatively consistent in their understanding of what is meant by individual intellectual capital. However, established definitions of its dimensions remain unspecific or lacking, and there is divergence among scholars regarding their interpretation. In particular, the definitions appear to leave the distinction between individual human capital and individual structural capital unclear. For example, Alppivuori and Kianto (2025) associate problem-solving skills and thinking methods with individual human capital, whereas Wang et al. (2020) classify them as part of individual structural capital.

2.3 Dimensions of individual intellectual capital: An organising framework

Because previous definitions of the individual dimensions of intellectual capital are either imprecise (Wang et al., 2020) or absent (Alppivuori and Kianto, 2025), and partly contradictory, there is a clear need for a proposal that both defines and specifies these dimensions. Such an elaboration is presented in Table 1. It draws in part on the earlier definitions of individual intellectual capital proposed by Wang et al. (2020), while further refining and specifying them. These definitions are subsequently employed as the organising framework for the review.

Table 1: Dimensions of individual intellectual capital

	Individual human capital	Individual structural capital	Individual relational capital
Definition	Knowledge and knowing capability embodied in an individual, encompassing knowledge, competencies, values, motivations, skills, and capacity for action.	Knowledge and knowing capability that are embedded in relatively stable knowledge structures, such as cognitive frames, mental models, habitual routines, methods, and professional practices, as well as more encompassing paradigms and worldviews.	Knowledge and knowing capability that arise from and are embedded in an individual's relationships, networks, and collaborative interactions.
Primary locus	Embodied in the individual	Embedded in relative stable knowledge structures that manifest as internalised patterns of thought and behaviour	Embedded in social relationships and co-created through interaction
Degree of stability	Dynamic and situational	Relatively stable and enduring over time	Dependent on quality and continuity of relationships

Drawing on the existing literature, this study conceptualises individual intellectual capital as the knowledge and knowing capacity embodied in an individual that can be mobilised to create value. Individual structural capital is understood to be relatively stable and enduring when compared with the more dynamic and context-dependent nature of individual human capital. As with organisational structural capital, changes to an individual's structural capital are generally more difficult to enact than changes to individual human capital.

Individual relational capital, by contrast, is embedded in an individual's social relationships and is co-created through ongoing interaction. Emotional and renewal capital, as identified by Alppivuori and Kianto (2025), are conceptualised as being subsumed within this tripartite framework, specifically within the dimensions of human and structural capital.

3 Methodology

This study employs the integrative review method (Torraco, 2005; 2016), which "reviews, critiques, and synthesises representative literature on a topic in an integrated manner such that new frameworks and perspectives on the topic are generated" (Torraco, 2005, p. 356). This approach is well suited to the aims of the study, as it facilitates the synthesis of perspectives and insights from multiple research fields and thereby supports the development of a novel conceptual framework of individual intellectual capital for sustainability. Integrative reviews often require a more creative approach to literature identification, as the researcher must weave together insights drawn from diverse bodies of research (Snyder, 2019). Accordingly, the study does not rely on systematic literature searches, but rather on the identification of relevant research domains from which seminal and foundational sources were sought.

Literature searches were conducted using the Scopus database and Google Scholar in order to identify highly cited sources within the following research domains: education for sustainable development—including transformative learning for sustainability and sustainability competences—sustainability leadership, sustainability transformation, and sustainability science. Highly cited articles were selected for analysis where they addressed forms of knowledge or knowing considered necessary for advancing sustainability. The final body of literature (n = 12) includes work on transformative learning for sustainability (Sipos, Battisti and Grimm, 2008; Sterling, 2011; Rodríguez Aboytes and Barth, 2020), sustainability competencies (Olsson et al., 2020; Brundiers et al., 2021; Redman and Wiek, 2021), sustainability leadership (Metcalf and Benn, 2013; Haney, Pope and Arden, 2020; Sajjad, Eweje and Raziq, 2024), as well as sustainability transformation and sustainability science (Amel et al., 2017; Abson et al., 2017; Wamsler et al., 2021).

Integrative literature reviews require conceptual structuring in order to clarify how the central concepts of a topic cohere into a unified whole (Torraco, 2005, 2016). In this study, the theory of individual intellectual capital is employed as a guiding theoretical framework to structure the topic conceptually and to organise the synthesis. Accordingly, the critical analysis and synthesis of key intangible assets for sustainability are undertaken by enriching the theory of individual intellectual capital with new insights, concepts, and theoretical perspectives drawn from the research fields of education for sustainable development, sustainability transformation, sustainability science, and sustainability leadership.

4 Sustainability Transformation through Individual Intellectual Capital

4.1 Individual human capital for sustainability

The role of critical, complexity, and systems thinking is widely recognised in the literature as a core sustainability competence (Metcalf and Benn, 2013; Rodríguez Aboytes and Barth, 2020; Brundiers et al., 2021; Redman and Wiek, 2021; Sajjad, Eweje and Raziq, 2024). The capacity to understand socio-ecological systems, engage with complexity and uncertainty, and reason through complex problems is regarded as crucial for advancing sustainability (Metcalf and Benn, 2013). In addition, futures thinking, values thinking, and strategic thinking are identified as important competencies for articulating sustainability visions, developing transition strategies, and applying, reconciling, and negotiating sustainability-related values.

However, the literature on individual human capital for sustainability does not focus solely on cognitive dimensions. The tripartite classification of learning domains into cognitive, social-emotional or affective, and behavioural or psychomotor domains is widely applied in education for sustainable development, and particularly within transformative learning theory for sustainability (Sipos et al., 2008; Rodríguez Aboytes and Barth, 2020). For example, Sipos et al. (2008) draw on the principles of 'Head' (engagement, cognitive), 'Hands' (enactment, psychomotor), and 'Heart' (enablement, affective) to facilitate transformative behavioural change towards sustainability. Accordingly, values and socio-emotional characteristics – such as empathy, compassion, collective concern, and non-materialistic and biospheric values – are emphasised as central components of sustainability learning (Rodríguez Aboytes and Barth, 2020; Wamsler et al., 2021). Sajjad, Eweje and Raziq (2024) argue that ethical values provide the foundation for sustainability leadership practice, while also highlighting the importance of mindfulness and resilience. Such personal wellbeing dimensions are emphasised in several other studies. Brundiers et al. (2021) and Redman and Wiek (2021), for instance, refer to intrapersonal competence, defined as the ability to maintain personal wellbeing and avoid adverse health impacts when engaging in sustainability-related work. Similarly, Metcalf and Benn (2013) stress the importance of emotional intelligence and the capacity to manage emotions.

In addition, a sense of agency, empowerment to act, and other action-oriented capabilities are highlighted across the literature (Haney, Pope and Arden, 2020; Olsson et al., 2020; Rodríguez Aboytes and Barth, 2020; Brundiers et al., 2021; Redman and Wiek, 2021). Olsson et al. (2020) conceptualise action competence for sustainability as comprising knowledge of action possibilities, confidence in one's own capacity to exert influence, and a willingness to act. Likewise, Brundiers et al. (2021) and Redman and Wiek (2021) identify implementation competence as a key sustainability competence, referring to the ability to put sustainability strategies into action in effective and efficient ways – that is, to engage in conscious, purposeful action.

4.2 Individual structural capital for sustainability

Many studies argue that sustainability transformations require changes at deeper levels, including shifts in worldviews and paradigms (Amel et al., 2017; Abson et al., 2017; Wamsler et al., 2021). Abson et al. (2017) point out to the need to change mindsets and ways of thinking, which are seen as the deep leverage points out of which the whole system arises (Abson et al. 2017). Amel et al. (2017) emphasise that leaders must think differently and internalise an ecologically grounded worldview. For example, in the context of food systems, such a worldview entails several assumptions that contrast with those of the modern-industrial worldview: Earth's regenerative capacity is limited (rather than abundant), socio-ecological systems must be circular (rather than linear "take-make-waste" systems that can continue indefinitely), even small manipulations of nature has system-wide consequences (rather than technology will fix our problems), and diversity maintains system integrity (rather than efficiency increases profits).

Alongside these more cognitively oriented worldview shifts, the literature highlights identity-based changes in individuals' relationships and connections with others, nature, and future generations (Amel et al., 2017; Abson et al., 2017; Wamsler et al., 2021). Haney, Pope and Arden (2020) emphasises the importance of personal connection to sustainability by allowing individuals to see and to feel the connections between themselves and particular sustainability issues, and between potential solutions to sustainability problems and their roles. Furthermore, Wamsler et al. (2021) emphasise agency as the capacity to recognise broader and deeper patterns and one's own role in the world, combined with the intention, optimism, and courage to act upon this understanding. Similarly, Amel et al. (2017) highlight the role of agency and individuals' influence on broader systems through ecologically compatible engagement and leadership. Individual structural capital for sustainability thus emerges across cognitive, identity-based, and agency-related dimensions.

Overall, individual structural capital appears to be emphasised through the knowledge and knowing capacities required to advance sustainability, as this perspective focuses on knowledge embedded in relatively

stable cognitive structures. Many authors stress the need to influence and transform these structures as deep leverage points for change (e.g. Sterling, 2011; Abson et al., 2017; Wamsler et al., 2021). The model of the levels of knowing by Sterling (2011) suggest that deeper perceptions and conceptions (related to metaphysics, paradigms and worldviews) inform, influence and help manifest beliefs, values, norms, and assumptions, which in turn influence more immediate ideas, theories and everyday thoughts and actions. Thus, individual structural capital forms the basis of individual intellectual capital, while intellectual capital manifests itself across different levels of knowing.

4.3 Individual relational capital for sustainability

Studies recognise the importance of interpersonal competence in advancing sustainability, as sustainability transformations require collaboration and the ability to engage others effectively (Amel et al., 2017; Rodríguez Aboytes and Barth, 2020; Brundiers et al., 2021; Redman and Wiek, 2021). Interpersonal competence refers to the capacity to work successfully in interdisciplinary and interprofessional teams, as well as to involve diverse stakeholders in meaningful and effective ways in advancing sustainability transformations (Redman and Wiek, 2021).

4.4 Individual intellectual capital for sustainability: A framework

Individual intellectual capital for sustainability is defined here as an asset encompassing the full range of individual cognitive, socio-emotional, and behavioural knowledge and knowing capabilities that contribute to the development of sustainability-oriented thinking, identity, and agency, and thereby support sustainability transformation. Table 3 illustrates the constituent elements of individual intellectual capital for sustainability. Each dimension of individual intellectual capital comprises the elements of thinking, identity, and agency, which together represent three interrelated domains of knowing.

Table 3. Individual intellectual capital for sustainability.

Individual intellectual capital for sustainability encompasses the cognitive, socio-emotional, and behavioural knowledge and knowing capabilities embodied in an individual that shape thinking, identity, and agency in ways that enable and advance sustainability transformation.			
	Individual human capital for sustainability	Individual structural capital for sustainability	Individual relational capital for sustainability
Sustainability thinking *(cognitive)*	Expertise and content knowledge of sustainability Systems, complex and critical thinking competencies, and reflexivity Futures, strategic and values thinking competencies	Ecologically grounded worldview	Relationships and networks to promote sustainability
Sustainability identity *(socio-emotional)*	Empathy, compassion, collective concerns, non-materialistic and biospheric values (ethical values) Intrapersonal competence, i.e., the ability to promote sustainability in self-caring ways, resilience, mindfulness and emotional intelligence	Interconnection with other people, society, and the more-than-human world Personal connection with sustainability	Collaborative capacity and will to promote sustainability in collaboration with others
Sustainability agency *(behavioural)*	Action and implementation competencies	Ecologically compatible engagement, agency and leadership	Interpersonal and integration competencies

Compared to conventional conceptualisations of individual intellectual capital, individual intellectual capital for sustainability foregrounds three interrelated aspects: a broad understanding of knowledge, the role of structural capital, and the importance of transformative competencies that enable change and action (for example, the ability to question and critique prevailing systems and institutions). A broad understanding of knowledge implies that individual human, structural, and relational capital each encompass the elements of

thinking, identity, and agency, rather than focusing exclusively on the cognitive aspects of knowledge. This broader conception also reflects the relevance of multiple levels of knowing, as emphasised by Sterling (2011).

The role of structural capital is particularly salient in the context of individual intellectual capital for sustainability. Previous studies have highlighted the fundamental importance of human capital. For instance, Wang, Bu and Cai (2021, p. 1206) argue that individual human capital constitutes the most important dimension, as it can generate the other two forms: "if individuals make efforts to study and learn objective knowledge, their professional skills and competence would increase, resulting in better individual structural capital and individual relational capital." However, when individual intellectual capital is considered from a sustainability perspective, the emphasis shifts away from the accumulation of objective knowledge towards the foundational role of structural capital. Ways of thinking and underlying worldviews are understood to form the basis for the meaningful application of more technical and specialised expertise in advancing sustainability.

Finally, the importance of transformative competencies is accentuated in conceptualisations of individual intellectual capital for sustainability. Intellectual capital for sustainability transformation is not limited to adopting efficient organisational routines; it also entails the capacity to challenge and transform existing practices, inspire others to engage in change processes, and reflectively learn from experience. Consequently, competencies such as systems thinking, critical thinking, strategic thinking, and futures thinking, along with reflection and implementation capacities, are particularly emphasised.

5 Conclusions

This study makes two contributions to intellectual capital theory. First, it conceptualises and defines individual intellectual capital for sustainability by synthesising insights from the literatures on sustainability leadership, sustainability science, sustainability competencies, and transformative learning for sustainability with intellectual capital theory. Individual intellectual capital for sustainability is defined as an asset encompassing the full range of individual cognitive, socio-emotional, and behavioural knowledge and knowing capabilities that contribute to the development of sustainability-oriented thinking, identity, and agency, and thereby promote sustainability transformation. In contrast to traditional definitions of intellectual capital, which primarily emphasise the creation of competitive advantage, this conceptualisation explicitly foregrounds the role of intellectual capital in enabling sustainability transformation. Research on individual intellectual capital remains limited. While sustainable intellectual capital have been conceptualised at the organisational level, they have not been systematically defined at the individual level. The proposed definition of individual intellectual capital for sustainability therefore addresses an important gap by emphasising a broader understanding of knowledge that extends beyond purely cognitive elements and immediate, concrete levels of knowing.

Second, the study specifies the content of individual intellectual capital for sustainability and develops a novel conceptual framework. Previous studies have operationalised individual intellectual capital through sets of measurable indicators (Wang et al., 2020; Alppivuori and Kianto, 2025), but without an explicit sustainability focus. Similarly, earlier work has conceptualised green intellectual capital in general terms (Chen, 2008), yet has not articulated its dimensions at the individual level. The framework presented in this study addresses these gaps by detailing the constituent elements of individual intellectual capital for sustainability.

The framework supports the management of human resources for sustainability transformation by identifying key intangible assets relevant to advancing sustainability. Moreover, the integrated framework offers a foundation for future research aimed at operationalising and measuring individual intellectual capital for sustainability. At present, there are few well-developed and widely accepted assessment methods for sustainability competencies or intellectual capital at the individual level. Research suggests that current practices for assessing students' sustainability competencies remain at an early stage of development (Redman et al., 2021), and that considerable ambiguity persists regarding the measurement scales employed (Annelin and Boström, 2023). Future studies could therefore focus on developing and validating robust measures of individual intellectual capital for sustainability, with potential benefits for both education and management research and practice.

From the perspective of education for sustainability, the intellectual capital lens highlights that sustainability competencies should be understood not only as learning outcomes, but also as manageable intangible assets that can be intentionally developed and mobilised to support sustainability transformation.

6 Ethics Declaration

The research referred to in this paper did not require ethical clearance, as no human subjects, personal data, or ethically sensitive materials were involved. Consequently, no formal ethical approval was sought or granted.

7 AI Declaration

Microsoft 365 Copilot Chat (Pro version with an organisational ID) was used to check the final text for grammatical errors. The tool was thus used for language improvement.

References

Abson, D. J., Fischer, J., Leventon, J., Newig, J., Schomerus, T., Vilsmaier, U., ... and Lang, D. J. (2017) "Leverage points for sustainability transformation", *Ambio*, Vol. 46, pp. 30-39.

Alppivuori, S. M. and Kianto, A. (2025) "Individual-Level Intellectual Capital: A Framework for Future Management and Reflection", In *Futurizing Intellectual Capital: Insights on Navigating Knowledge-Based Value Creation* (pp. 65-91), Cham: Springer Nature Switzerland.

Alvino, F., Di Vaio, A., Hassan, R. and Palladino, R. (2021) "Intellectual capital and sustainable development: A systematic literature review", *Journal of Intellectual Capital*, Vol. 22 No. 1, pp. 76-94.

Amel, E., Manning, C., Scott, B., and Koger, S. (2017) "Beyond the roots of human inaction: Fostering collective effort toward ecosystem conservation", *Science*, Vol. 356 No. 6335, pp. 275-279.

Annelin, A., and Boström, G. O. (2023) "An assessment of key sustainability competencies: a review of scales and propositions for validation", *International Journal of Sustainability in Higher Education*, Vol. 24 No. 9, pp. 53-69.

Arenas, T., Griffiths, P., and Freraut, A. (2013) "An individual-centred model of Intellectual Capital", In Proceedings of the International Conference on Intellectual Capital, Knowledge Management and Organizational Learning, Vol. 8 No. 7, pp. 46-53.

Benevene, P., Buonomo, I., Kong, E., Pansini, M. and Farnese, M. L. (2021) "Management of green intellectual capital: Evidence-based literature review and future directions", *Sustainability*, Vol. 13 No. 15, 8349.

Brundiers, K., Barth, M., Cebrián, G., Cohen, M., Diaz, L., Doucette-Remington, S., ... and Zint, M. (2021), "Key competencies in sustainability in higher education—toward an agreed-upon reference framework", *Sustainability Science*, Vol. 16, pp. 13-29.

Čabrilo, S., Užienė, L. and Kianto, A. (2025) "Rethinking the future of intellectual capital: emerging perspectives in IC theory and practice", In ECKM 2025: proceedings of the 25th European conference on knowledge management, 4-5 September, 2025, Lahti, Finland. (Vol. 26, No. 1, pp. 166-172). ACPI.

Chen, Y.S. (2008) "The Positive Effect of Green Intellectual Capital on Competitive Advantages of Firms", *Journal of Business Ethics*, Vol. 77, pp. 271-286.

Haney, A. B., Pope, J. and Arden, Z. (2020), "Making it personal: Developing sustainability leaders in business", *Organization & Environment*, Vol. 33 No. 2, pp. 155-174.

Hong, Y., Chen, T., & Zhang, Y. (2025) Individual intellectual capital, creative process engagement and information security policy compliance. *Journal of Intellectual Capital*, 26(6), 1233-1256.

Kang, S. C. and Snell, S. A. (2009) "Intellectual capital architectures and ambidextrous learning: a framework for human resource management", *Journal of Management Studies*, Vol. 46 No. 1, pp. 65-92.

Khan, M. M., Mubarik, M. S., Ahmed, S. S., Islam, T., and Rehman, S. U. (2024) "Utilizing every grain of intellect: exploring the role of individual-level intellectual capital in linking servant leadership with innovative work behavior", *Journal of Intellectual Capital*, Vol. 25 No. 1, pp. 23-37.

Kmieciak, R. (2024) "Do employees with high human capital hide knowledge? Exploring mediation and moderation mechanisms", *Journal of Intellectual Capital*, Vol. 25, No. 5/6, pp. 1042-1061.

Martín-de-Castro, G., Delgado-Verde, M., López-Sáez, P. and Navas-López, J. E. (2011), "Towards 'an intellectual capital-based view of the firm': origins and nature", *Journal of Business Ethics*, Vol. 98, pp. 649-662.

Metcalf, L., & Benn, S. (2013) "Leadership for sustainability: An evolution of leadership ability", *Journal of Business Ethics*, 112(3), 369-384.

Nahapiet, J., and Ghoshal, S. (1998) "Social capital, intellectual capital, and the organizational advantage", *Academy of Management Review*, Vol. 23, No. 2, pp. 242-266.

Olsson, D., Gericke, N., Sass, W., & Boeve-de Pauw, J. (2020) "Self-perceived action competence for sustainability: The theoretical grounding and empirical validation of a novel research instrument", *Environmental Education Research*, 26(5), 742-760.

Redman, A. and Wiek, A. (2021) "Competencies for advancing transformations towards sustainability", *Frontiers in Education*, Vol. 6, p. 785163.

Redman, A., Wiek, A. and Barth, M. (2021) "Current practice of assessing students' sustainability competencies: A review of tools", *Sustainability Science*, Vol. 16, pp. 117-135.

Richardson, K., Steffen, W., Lucht, W., Bendtsen, J., Cornell, S. E., Donges, J. F., ... and Rockström, J. (2023) "Earth beyond six of nine planetary boundaries", *Science Advances*, Vol. 9 No. 37, eadh2458.

Rodríguez Aboytes, J. G., and Barth, M. (2020) "Transformative learning in the field of sustainability: a systematic literature review (1999-2019)", *International Journal of Sustainability in Higher Education*, Vol. 21 No. 5, pp. 993–1013.

Sajjad, A., Eweje, G., and Raziq, M. M. (2024) "Sustainability leadership: An integrative review and conceptual synthesis", *Business Strategy and the Environment*, Vol. 33 No. 4, pp. 2849-2867.

Sipos, Y., Battisti, B. and Grimm, K. (2008) "Achieving transformative sustainability learning: engaging head, hands and heart", *International Journal of Sustainability in Higher Education*, Vol. 9 No. 1, pp. 68-86.

Snyder, H. (2019) Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, 104, 333-339.

Sterling, S. (2011) "Transformative learning and sustainability: Sketching the conceptual ground", *Learning and Teaching in Higher Education*, Vol. 5 No. 11, pp. 17-33.

Stockholm Resilience Centre (2026) "Planetary boundaries", available at: https://www.stockholmresilience.org/research/planetary-boundaries.html (accessed 27 April 2026).

Subramaniam, M. and Youndt, M. A. (2005) "The influence of intellectual capital on the types of innovative capabilities", *Academy of Management Journal*, Vol. 48 No. 3, pp. 450-463. https://doi.org/10.5465/amj.2005.17407911

Torraco, R. J. (2005) "Writing integrative literature reviews: Guidelines and examples". *Human Resource Development Review*, Vol. 4 No. 3, pp. 356-367.

Torraco, R. J. (2016) "Writing integrative literature reviews: Using the past and present to explore the future", *Human Resource Development Review*, Vol. 15, No. 4, pp. 404-428.

Wamsler, C., Osberg, G., Osika, W., Herndersson, H., & Mundaca, L. (2021) "Linking internal and external transformation for sustainability and climate action: Towards a new research and policy agenda", *Global Environmental Change*, 71, 102373.

Wang, Z., Cai, S., Liu, M., Liu, D., and Meng, L. (2020) "The effects of self-reflection on individual intellectual capital", *Journal of Intellectual Capital*, Vol. 21 No. 6, pp. 1107-1124.

Wang, Z., Bu, X. and Cai, S. (2021) "Core self-evaluation, individual intellectual capital and employee creativity", *Current Psychology*, Vol. 40 No. 3, pp. 1203-1217.

Yusliza, M. Y., Yong, J. Y., Tanveer, M. I., Ramayah, T., Faezah, J. N. and Muhammad, Z. (2020) "A structural model of the impact of green intellectual capital on sustainable performance", *Journal of Cleaner Production*, Vol. 249, 119334.

Dynamic Capabilities and Economies of Scale: An Optimization Perspective on Intelligent Knowledge Investments

László Kállay
Corvinus University of Budapest

Abstract

This paper develops an economic perspective on dynamic capabilities by conceptualizing them as investments in intelligent knowledge under uncertainty. The dynamic capabilities literature has made substantial progress in explaining how firms sense opportunities, seize them, and reconfigure resources, yet it has paid less systematic attention to the cost structure of capability development and to the possibility that organizations may invest either too little or too much in dynamic capabilities. Building on a formal transaction cost optimization model previously developed by the author, the paper argues that dynamic capability investment follows an analogous optimization logic: organizations incur direct and quasi-fixed costs in building sensing, seizing, and reconfiguring mechanisms, while these investments reduce the expected losses associated with misadaptation, missed opportunities, strategic errors, or delayed transformation. The central claim is that there is a scale-dependent optimum of dynamic capability investment. As organizational exposure increases - through size, volume, scope, technological interdependence, or ecosystem involvement - the potential losses from weak adaptation increase, making higher capability investment economically justified. However, the optimum remains bounded because excessive investment may generate coordination burdens, cognitive overload, bureaucratic routines, and declining flexibility. The paper contributes to dynamic capability theory by integrating economies and diseconomies of scale into the analysis of intelligent knowledge investments and by deriving propositions for SMEs, large firms, and sustainability-oriented innovation ecosystems.

Keywords: dynamic capabilities; economies of scale; intelligent knowledge; transaction costs; optimization; sustainability-oriented innovation

Paper type: Academic Research Paper

1 Introduction

Dynamic capabilities have become one of the central concepts of strategic management because they explain how organizations renew themselves under conditions of uncertainty, technological change, and competitive turbulence. Since Teece, Pisano, and Shuen (1997), dynamic capabilities have been understood as higher-order capacities that enable firms to integrate, build, and reconfigure internal and external competences. Later work refined this broad definition by distinguishing sensing, seizing, and reconfiguring as core classes of capability activity (Teece, 2007), while the microfoundations literature emphasized managerial cognition, routines, decision structures, governance mechanisms, and organizational learning as the building blocks of such capabilities.

A large part of this literature is concerned with why some firms adapt better than others. It explains differences in strategic renewal, innovation performance, and resilience by reference to heterogeneous capabilities accumulated through path-dependent learning. This is a powerful perspective, but it also leaves an important question underdeveloped: how much should an organization invest in dynamic capabilities? The literature often treats dynamic capabilities as desirable strategic assets, but it rarely treats them as costly investments with an optimum level. This omission matters because dynamic capabilities are not costless. They require environmental scanning, data systems, managerial attention, strategic forums, experimentation budgets, cross-functional coordination, change management routines, and governance mechanisms. These investments may improve adaptation, but they also consume resources and may themselves become a source of complexity.

This paper addresses this gap by developing an optimization perspective on dynamic capability investment. The argument is not that dynamic capabilities should be reduced to a simple cost item. Rather, the paper claims that the development and deployment of dynamic capabilities has a recognizable economic structure. Capability investments create direct and quasi-fixed costs, while reducing expected losses from uncertainty, misadaptation, delayed response, or missed opportunities. The relevant managerial problem is therefore not whether dynamic capabilities are valuable in general, but which level and architecture of dynamic capabilities are justified under given conditions of organizational scale, scope, and exposure.

The paper builds on a formal transaction cost optimization model previously developed by the author (Kállay, Takács, & Trautmann 2020). That model demonstrates that total cost under uncertainty can be represented as the sum of direct costs and expected losses, and that an optimum exists because higher direct expenditure reduces the probability of loss but with diminishing returns. The present paper adapts this logic to dynamic capabilities. In the proposed framework, investment in sensing, seizing, and reconfiguring reduces the expected loss associated with weak adaptation. However, as investment rises, marginal returns decline and coordination costs increase. The result is a scale-dependent optimum.

The main contribution of the paper is threefold. First, it integrates economies and diseconomies of scale into dynamic capability theory. Second, it conceptualizes intelligent knowledge as an economically constrained organizational infrastructure rather than merely as a knowledge stock. Third, it develops propositions linking organizational scale to optimal dynamic capability investment and to the risk of underinvestment or overinvestment. The framework is particularly relevant for sustainability-oriented innovation, where organizations must combine experimentation, adaptation, and diffusion at scale.

2 Literature Background

2.1 Dynamic capabilities and intelligent knowledge

Dynamic capabilities research starts from the observation that ordinary operational efficiency is insufficient when environments change. Operational capabilities allow firms to perform current activities effectively; dynamic capabilities allow them to modify the resource base and the activity system itself (Winter, 2003; Helfat & Winter, 2011). Teece (2007) distinguishes three broad groups of microfoundations: sensing opportunities and threats, seizing opportunities through investment and business model choices, and reconfiguring assets and organizational structures. These activities are knowledge-intensive. They depend on how organizations collect information, interpret signals, coordinate judgment, and implement change.

This is naturally linked to the notion of intelligent knowledge. In this paper, intelligent knowledge is not treated as a passive stock of information. It is defined as the capability-embedded ability of an organization to generate, integrate, interpret, and apply knowledge for purposeful adaptation. It is intelligent because it connects information to action; it is organizational because it is embedded in routines, governance structures, and decision processes; and it is dynamic because its value is realized under uncertainty, not merely under stable execution conditions.

The dynamic capabilities literature has strongly emphasized heterogeneity, path dependence, and context specificity. These are essential insights. However, the same emphasis has also made the literature less attentive to the cost and scalability of capability development. If capabilities are highly idiosyncratic and embedded, then replication and scaling are difficult. Yet many firms do in fact create shared innovation platforms, central analytics units, corporate venture functions, transformation offices, and standardized strategic review processes. These are attempts to make at least some parts of dynamic capability activity scalable.

2.2 Economies of scale in capability development

Economies of scale arise when average costs decline as output, volume, or scope increases. In the context of dynamic capabilities, the relevant output is not a physical product but the number and value of adaptive decisions, innovation initiatives, transformations, and strategic reallocations supported by the capability system. Many elements of dynamic capabilities have fixed or quasi-fixed costs. Examples include data infrastructure, strategic scanning systems, innovation governance, experimentation platforms, business intelligence routines, and specialized managerial expertise. Once created, these elements may serve multiple products, markets, business units, or ecosystem relationships.

The resource-based and capability literatures already contain partial support for this view. Penrose (1959) emphasized the role of managerial resources and the limits of firm growth, while also showing that unused productive services of resources may create opportunities for expansion. Zollo and Winter (2002) show that deliberate learning, articulation, and codification contribute to the evolution of capabilities. Winter and Szulanski (2001) discuss replication as a strategy, emphasizing that templates can reduce the cost of reproducing complex routines. These contributions imply that capability development can benefit from learning, codification, and reuse.

At the same time, dynamic capabilities are not automatically scalable. Szulanski (1996) shows that best practices are often sticky and difficult to transfer. Eisenhardt and Martin (2000) argue that dynamic capabilities differ across environmental conditions: in moderately dynamic markets they may be more structured and

analyzable, whereas in high-velocity environments they are more experiential and fragile. This implies that scalability depends on the balance between codification and contextual adaptation.

2.3 Diseconomies of scale and organizational limits

Scale also creates diseconomies. As organizations grow, coordination becomes more complex, communication channels multiply, and decision processes slow down. A larger organization may have more resources for sensing and seizing, but it may also suffer from internal fragmentation and cognitive overload. Helfat and Winter (2011) caution that the distinction between operational and dynamic capabilities is not always clear in practice, and that capabilities must be understood in relation to their actual use. More dynamic capability activity is not automatically better if it creates excessive deliberation, redundant committees, or rigid change programs.

The key theoretical implication is that dynamic capability investment should not be assumed to have monotonically positive returns. At low levels, investment may sharply reduce the risk of misadaptation. At higher levels, additional investment may yield smaller incremental reductions in expected loss while adding direct cost and complexity. The natural implication is an interior optimum: a level of investment at which the marginal cost of further capability development equals the marginal reduction in expected loss.

3 Conceptual Framework: Dynamic Capability Investment as Optimization

3.1 From transaction cost optimization to capability investment

The model used as a foundation in this paper defines total cost as the sum of direct costs and expected losses. In a transaction, actors may spend resources to reduce the probability of loss. The expected loss is the probability of loss multiplied by exposure. The total cost is minimized where the saving in expected loss from an additional unit of effort is equal to the additional direct cost. The model also shows that the optimum shifts when exposure changes: at low exposure, little effort may be optimal; at high exposure, more effort is justified.

This logic can be transferred to dynamic capabilities without claiming that dynamic capabilities are identical to transaction costs. The structural analogy is the important point. Organizations invest in dynamic capabilities to reduce the probability and consequences of strategic failure under uncertainty. Weak sensing may lead to missed opportunities. Weak seizing may lead to delayed or mistaken investment. Weak reconfiguring may lead to inertia, stranded assets, or failed transformation. These are expected losses. Dynamic capability investment reduces them, but only by incurring direct and quasi-fixed costs.

Table 1. Mapping the transaction cost model to dynamic capability investment

Transaction cost model	Dynamic capability interpretation	Illustrative examples
Direct cost of effort	Investment in sensing, seizing, and reconfiguring	Analytics systems, strategic forums, experimentation budgets, transformation offices
Probability of loss	Probability of misadaptation or missed opportunity	Wrong market entry, delayed innovation, failure to respond to regulation
Exposure	Organizational scale, scope, volume, or strategic stakes	Number of products, markets, transactions, customers, ecosystem links
Expected loss	Value at risk from weak adaptation	Lost revenue, technological lock-in, reputational damage, compliance risk
Cost sharing rule	Internal or inter-organizational allocation of capability costs and benefits	Business unit budgets, ecosystem governance, platform participation rules

Source: Author, building on (Kállay, Takács, & Trautmann 2020)) and the dynamic capabilities literature.

3.2 Formal representation

Let I denote the level of investment in dynamic capabilities and S denote organizational scale or exposure. Scale may refer to firm size, activity volume, product-market scope, technological complexity, or the value at stake in adaptation decisions. The organization chooses I to minimize total adaptation cost:

$$TC(I,S) = C(I) + L(I,S)$$

where C(I) is the direct and quasi-fixed cost of capability development, and L(I,S) is the expected loss associated with uncertainty and misadaptation. The cost function C(I) is increasing in I. The expected loss L(I,S) is decreasing in I, because stronger dynamic capabilities reduce the probability or magnitude of failure, but it is increasing in S, because larger organizations or broader activity systems face greater exposure.

The optimality condition is:

dC(I)/dI = - dL(I,S)/dI

This condition states that the organization should invest in dynamic capabilities up to the point where the marginal cost of additional capability development equals the marginal reduction in expected loss. If investment is lower than this point, the organization leaves avoidable losses unaddressed. If investment is higher, the organization spends more on capability infrastructure than the expected reduction in losses justifies.

3.3 Scale-dependent optimum

Figure 1 illustrates the central mechanism. Direct capability costs increase with investment, while expected losses decline as stronger dynamic capabilities reduce misadaptation. Total cost is therefore minimized at an interior optimum. When scale or exposure rises, the total cost curve shifts and the optimum moves to the right, implying a higher economically justified level of dynamic capability investment.

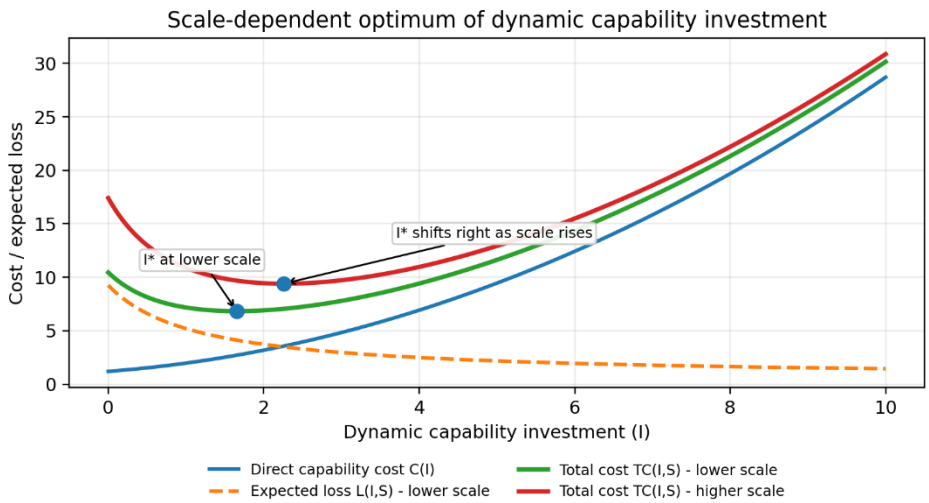


Figure 1. Scale-dependent optimum of dynamic capability investment. Source: Author.

4 Economies and Diseconomies of Scale in Dynamic Capabilities

4.1 Economies in development

Economies of scale in dynamic capability development arise when the average cost of capability infrastructure declines as the organization uses it across more decisions, projects, products, or units. A central market intelligence function may be expensive to create, but once it exists, it can inform many strategic choices. A data platform may require substantial investment, but its marginal cost of supporting an additional innovation project may be low. A corporate venture process or transformation office may be costly, but its templates and routines may be reused across multiple initiatives.

Learning economies are also central. Repeated engagement in innovation and transformation helps organizations learn how to search, experiment, evaluate, and reconfigure. Zollo and Winter (2002) emphasize deliberate learning mechanisms such as experience accumulation, knowledge articulation, and codification. These mechanisms reduce the cost of future capability deployment by transforming repeated experience into organizational routines. In this sense, dynamic capabilities become more efficient when they are used often enough to generate learning.

Replication economies are related but distinct. Winter and Szulanski (2001) show that replication requires a template and a process for transferring routines. Dynamic capabilities may be scalable when sensing dashboards, innovation stage-gates, portfolio review practices, or transformation playbooks can be transferred across units. However, replication is never purely mechanical. The local context matters, and excessive standardization may undermine the adaptive purpose of the capability.

4.2 Economies in deployment

Capability development and capability deployment should be distinguished. Development refers to building the knowledge infrastructure, routines, and governance mechanisms that make dynamic adaptation possible.

Deployment refers to applying these mechanisms in concrete contexts. Economies of scale may be stronger in development than in deployment. It may be efficient to build a common data infrastructure or strategic review process, but the interpretation of signals and the reconfiguration of resources may still require local judgment.

This distinction helps explain why large firms often develop formal systems for innovation and transformation, while still struggling to achieve agility. The common infrastructure creates economies, but deployment across heterogeneous business units may produce transfer frictions. Szulanski (1996) concept of internal stickiness is useful here: knowledge does not move automatically just because it is available. Motivation, absorptive capacity, causal ambiguity, and relationship quality affect transfer.

4.3 Diseconomies and the limits of capability investment

Diseconomies of scale appear when additional investment in dynamic capabilities creates more complexity than adaptive benefit. More scanning may generate information overload. More governance may slow down decisions. More formalization may reduce entrepreneurial initiative. More coordination may increase political bargaining inside the firm. These effects are especially relevant for large organizations, where dynamic capability systems can become bureaucratic.

The diseconomy argument is not a rejection of dynamic capabilities. It is a warning against treating them as unbounded goods. The same routines that support adaptation at one level may produce rigidity at another. A transformation office may accelerate resource reconfiguration when it clarifies priorities, but may hinder adaptation when it becomes a procedural bottleneck. A strategic planning system may help sense trends, but may create false confidence if it replaces experimentation with reporting.

5 Propositions

Proposition 1 - Optimization structure: Dynamic capability investment follows an optimization logic in which organizations balance the direct and quasi-fixed costs of capability development against the expected losses associated with uncertainty, misadaptation, and missed opportunities.

Proposition 2 - Scale-dependent optimum: The optimal level of dynamic capability investment increases with organizational scale, scope, and exposure, because larger or more complex organizations face higher expected losses from weak adaptation.

Proposition 3 - Bounded optimum: The optimal level of dynamic capability investment is finite and bounded because marginal returns decline and coordination costs increase at higher levels of investment.

Proposition 4 - Development-deployment asymmetry: Economies of scale are stronger in the development of dynamic capability infrastructure than in the deployment of dynamic capabilities across heterogeneous contexts.

Proposition 5 - Governance and cost allocation: The realized level of dynamic capability investment depends on how costs and benefits are allocated inside firms and across ecosystems. Misaligned internal cost-sharing rules may generate underinvestment, overinvestment, or fragmented investment.

Proposition 6 - Ecosystem scale effects: For sustainability-oriented innovation, the relevant scale of dynamic capability investment may be ecosystem-level rather than firm-level, because sensing, seizing, and reconfiguring often require inter-organizational coordination.

6 Illustrative Examples and Empirical Implications

6.1 SMEs

Small and medium-sized enterprises often face a structural disadvantage in dynamic capability investment. They need sensing, seizing, and reconfiguring capabilities no less than larger firms, but they have fewer opportunities to spread fixed costs. A professional market intelligence unit, data analytics system, or formal transformation office may be too expensive relative to the scale of the firm. Therefore, SMEs often rely on simpler, more informal, and more owner-manager centered forms of dynamic capability.

This does not mean that SMEs lack dynamic capabilities. Many SMEs are highly adaptive precisely because they are close to customers, have short communication lines, and can make quick decisions. Their problem is different: they may be unable to finance specialized knowledge infrastructure or to codify routines in a way that supports repeated scaling. The model predicts that the optimal level of formal DC investment is lower for SMEs, but also that underinvestment may occur when exposure grows faster than capability infrastructure.

6.2 Large firms and multinational corporations

Large firms can justify more extensive dynamic capability investments because the value at stake is higher. A mistake in technology strategy, regulatory response, sustainability transition, or platform positioning may affect many products and markets. Large firms can spread the cost of sensing systems, strategic analytics, corporate venture units, and reconfiguration routines across a broad activity base. This is the scale advantage emphasized by the model.

Yet large firms also face stronger diseconomies. Their dynamic capability systems may become too formal, too slow, or too politically mediated. Investment may continue beyond the optimum because managers prefer visible capability infrastructure, consultants, dashboards, and committees even when marginal adaptive value is low. Therefore, large firms face a double risk: they need more dynamic capability investment than SMEs, but they are also more exposed to overinvestment and bureaucratization.

6.3 Platform firms and digital ecosystems

Platform firms provide a clear example of scalable dynamic capability infrastructure. Digital platforms can use shared data, analytics, experimentation systems, and ecosystem governance to sense changes across many users and complementors. Once built, these systems may be deployed repeatedly at low marginal cost. A/B testing, automated monitoring, and algorithmic recommendation systems are examples of capability elements that can scale across very large volumes.

However, platform scale also increases exposure. A design error, governance failure, or regulatory misjudgment can have systemic consequences. Therefore, platform firms may require higher investment in sensing and governance than traditional firms, but they also face the risk that control systems become too centralized or insensitive to local ecosystem effects.

6.4 Sustainability-oriented innovation ecosystems

Sustainability transitions create especially strong reasons to analyze dynamic capabilities at the ecosystem level. Decarbonization, circular economy models, and inclusive innovation require coordinated changes among firms, public agencies, users, financiers, and knowledge institutions. No single organization fully controls the relevant resource base. Sensing must include technological, social, environmental, and regulatory signals. Seizing often requires joint investment. Reconfiguring may involve supply chains, standards, infrastructures, and institutions.

In such contexts, the optimal level of dynamic capability investment may not be determined at the level of the individual firm. A small firm may not be able to build advanced sensing capabilities alone, but a cluster, platform, or public-private partnership may create shared capability infrastructure. This creates collective economies of scale. At the same time, cost-sharing and governance problems become central. If some actors pay for shared sensing while others capture most of the benefits, investment may fall below the system-level optimum.

Table 2. Expected investment patterns by organizational context

Context	Scale effect	Main risk	Likely capability architecture
SMEs	Limited cost spreading	Underinvestment when exposure grows	Informal routines, owner-manager sensing, external networks
Large firms	Strong fixed-cost spreading	Overinvestment and bureaucracy	Formal analytics, governance, portfolio routines
Platform firms	Very strong deployment economies	Systemic failure and centralization	Data-driven sensing, scalable experimentation, ecosystem rules
Sustainability ecosystems	Collective economies beyond firm boundaries	Free-riding and weak cost sharing	Shared knowledge infrastructure, public-private governance

Source: Author.

7 Discussion

7.1 Contribution to dynamic capability theory

The paper contributes to dynamic capability theory by shifting attention from the existence of capabilities to their economically justified level. This is not a minor change. If dynamic capabilities are costly investments, then their value depends not only on whether they exist, but also on whether their scale, architecture, and

governance match the organizational exposure. The concept of a scale-dependent optimum helps reconcile two strands of the literature: the view that dynamic capabilities are necessary for adaptation and the view that organizational complexity can undermine responsiveness.

The proposed framework also clarifies why empirical findings on dynamic capabilities and performance may be mixed. If the relationship is non-linear, then studies assuming a simple positive association may miss important effects. Firms with low dynamic capability investment may perform poorly because they fail to adapt. Firms with excessive investment may also perform poorly because they suffer from complexity and slow decision-making. The highest performance should be expected near the optimum, not necessarily at the highest observed level of capability investment.

7.2 Contribution to IFKAD: intelligent knowledge as investment

For the IFKAD theme, the paper offers a specific interpretation of intelligent knowledge. Intelligent knowledge is not only knowledge that is useful, actionable, or technology-enabled. It is also knowledge embedded in costly organizational mechanisms. Its value depends on whether the organization can generate, process, and apply it at an appropriate scale. Too little intelligent knowledge infrastructure leaves the organization exposed to uncertainty. Too much may create reporting burdens, attention dispersion, and procedural inertia.

This interpretation connects knowledge management with strategic economics. It suggests that organizations should ask not only how to create more knowledge, but also where knowledge investments reduce expected losses most effectively. In sustainability-oriented organizations, this question becomes especially important because the relevant uncertainty is multidimensional: technological, regulatory, social, environmental, and reputational.

7.3 Managerial implications

Managers should treat dynamic capabilities as investments with explicit cost-benefit logic. This requires identifying the main organizational exposure: the value at risk from weak adaptation. For a small local firm, exposure may be modest and highly formal capability systems may be unnecessary. For a multinational firm, exposure may be large and underinvestment in sensing or reconfiguration may be extremely costly. For an ecosystem, exposure may be distributed, requiring shared governance and cost-sharing arrangements.

Managers should also distinguish between capability infrastructure and capability use. Building dashboards, offices, committees, and routines is not the same as improving adaptation. The relevant test is whether these investments reduce the probability or expected cost of strategic error. Finally, organizations should periodically reassess the optimum because both exposure and technology change. Digital tools may reduce the cost of sensing, while regulatory or sustainability pressures may increase the expected loss from weak adaptation.

8 Methodological Implications and Future Research

The framework suggests several empirical research paths. First, survey-based studies could measure dynamic capability investment, organizational scale, and performance outcomes, while testing for non-linear relationships. If the model is correct, the effect of investment should be positive up to a threshold and weaker or negative beyond it. Second, case studies could examine how firms decide on the size of sensing, innovation, and transformation functions. Such studies could identify whether managers explicitly consider exposure and expected loss or whether investment is driven by imitation and managerial fashion.

Third, comparative studies of SMEs and large firms could examine how fixed costs shape capability architecture. SMEs may substitute external networks and informal routines for internal systems, while large firms may rely on formalized processes. Fourth, ecosystem-level research could analyze cost-sharing rules in sustainability initiatives, clusters, and platforms. The transaction cost model suggests that inefficient cost-sharing may prevent the system from reaching the optimum even when the technical solution is known.

A useful empirical strategy would combine qualitative and quantitative elements. Qualitative case studies can identify the relevant capability investments and expected loss categories. Quantitative analysis can then examine whether investment intensity varies systematically with scale and whether performance follows an inverted-U pattern. Data envelopment analysis may also be relevant, because the original transaction cost model uses frontier logic to identify efficient combinations of direct costs and loss reduction. A similar approach could compare organizational units or firms as decision-making units using inputs such as DC investment and outputs such as innovation success, adaptation speed, or resilience.

9 Limitations

The paper is conceptual and does not provide direct empirical testing. Its purpose is to develop a theoretical framework that can guide subsequent empirical research. The formal representation is deliberately simple. It abstracts from multiple capability dimensions, uncertainty about the cost and benefit functions, and political processes inside organizations. These simplifications are useful for theory building, but future work should relax them.

Another limitation is that dynamic capabilities are difficult to measure. Investment in dynamic capabilities is not always visible in accounting data. It may be embedded in managerial attention, routines, culture, and informal networks. Therefore, empirical testing requires careful operationalization. Finally, the framework assumes that organizations can approximate the relationship between investment, expected loss, and scale. In practice, managers face imperfect information, and misperception of exposure may itself be a source of underinvestment or overinvestment.

10 Conclusion

This paper has developed an optimization perspective on dynamic capabilities and economies of scale. It argues that dynamic capabilities should be understood as costly investments in intelligent knowledge that reduce expected losses under uncertainty. Building on a formal transaction cost optimization model, the paper shows that the optimal level of dynamic capability investment depends on organizational scale or exposure. Larger organizations and broader ecosystems can justify higher investment because the value at risk from weak adaptation is larger. However, investment remains bounded because excessive capability infrastructure may generate coordination burdens, cognitive overload, and rigidity.

The paper contributes by integrating scale effects into dynamic capability theory, by conceptualizing intelligent knowledge as an economically constrained organizational infrastructure, and by deriving propositions for SMEs, large firms, platform organizations, and sustainability-oriented ecosystems. The central managerial implication is straightforward: organizations should not ask whether dynamic capabilities are good in general. They should ask what level, form, and governance of dynamic capability investment is optimal for their scale, exposure, and strategic uncertainty.

For future research, the framework opens a path toward empirical testing of non-linear capability-performance relationships and toward deeper analysis of cost-sharing rules in ecosystems. For practitioners, it provides a disciplined way to think about capability development: invest enough to reduce expected losses from uncertainty, but avoid building capability systems whose complexity exceeds their adaptive value.

11 Ethics Declaration

Ethical clearance was not required for the research.

12 AI Declaration

AI was used to improve the wording of this paper and finalizing the list of references.

References

Collis, D. J. (2019). The value of competitive advantage. Harvard Business School Working Paper No. 20-029.

Eisenhardt, K. M., & Martin, J. A. (2000). Dynamic capabilities: What are they? Strategic Management Journal, 21(10-11), 1105-1121.

Helfat, C. E. (2009). Understanding dynamic capabilities: Progress along a developmental path. Strategic Organization, 7(1), 91-102.

Helfat, C. E., & Winter, S. G. (2011). Untangling dynamic and operational capabilities: Strategy for the (n)ever-changing world. Strategic Management Journal, 32(11), 1243-1250.

Kállay, L., Takács, T. & Trautmann, L., (2020). Transaction costs: Economies of scale, optimum, equilibrium and efficiency - A game theory-based model of transaction costs. Preprints, 202010.0535.v1.

Penrose, E. T. (1959). The theory of the growth of the firm. Oxford University Press.

Szulanski, G. (1996). Exploring internal stickiness: Impediments to the transfer of best practice within the firm. Strategic Management Journal, 17(S2), 27-43.

Teece, D. J. (2007). Explicating dynamic capabilities: The nature and microfoundations of sustainable enterprise performance. Strategic Management Journal, 28(13), 1319-1350.

Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic capabilities and strategic management. Strategic Management Journal, 18(7), 509-533.

Wenzel, M., Danner-Schröder, A., & Spee, A. P. (2021). Dynamic capabilities? Unleash their dynamics through routine dynamics. Organization Studies, 42(1), 1-28.
Winter, S. G. (2003). Understanding dynamic capabilities. Strategic Management Journal, 24(10), 991-995.
Winter, S. G., & Szulanski, G. (2001). Replication as strategy. Organization Science, 12(6), 730-743.
Zollo, M., & Winter, S. G. (2002). Deliberate learning and the evolution of dynamic capabilities. Organization Science, 13(3), 339-351.

Do Renewable Energy Communities Create Value? Evidence from a Systematic Literature Review

Fabrizio Rossi, Alessandro Sicali
Università degli studi di Trieste, Trieste, Italy

Abstract

This study aims to analyse the literature on the economic performance of Renewable Energy Communities (RECs). Specifically, the objective of this work is to identify how existing research assesses REC profitability and to shed light on whether they represent financially viable investment vehicles. The distribution mechanisms employed to allocate economic benefits among REC participants are also investigated. To address these questions, a systematic literature review was conducted following the PRISMA 2020 guidelines, based on a search of the Scopus database. A final sample of 40 peer-reviewed articles published between 2018 and 2026 was analysed. The findings show that capital budgeting tools, such as NPV, IRR, and Payback Period, are the most widely adopted instruments for assessing REC economic performance. The empirical evidence suggests that RECs generally represent profitable investments capable of generating positive returns. Italy emerges as the most extensively studied context, notably in the residential sector, driven by a robust incentive system that has proven decisive in sustaining profitability. Indeed, economic performance tends to deteriorate and NPV can turn negative where regulatory support is insufficiently developed. Furthermore, photovoltaic technology emerges as the primary renewable energy source, while alternative configurations show potential but remain financially constrained by high CAPEX. Public and private involvement consistently generates positive economic outcomes by leveraging the complementarity of energy demand profiles. Regarding benefit distribution, a delicate balance exists between prosumers and consumers that can lead to participant dropout when allocation mechanisms prove ineffective. Minimum return guarantees, fairness-based criteria, and technology-driven tools have been shown to improve distributional outcomes and community stability. This study adds to the existing literature by providing systematic evidence on the key drivers of REC profitability and benefit distribution, while stimulating future research to further examine RECs through a corporate finance lens. The insights suggest that participation by private citizens, SMEs, and public administrations represents both an economically sustainable choice and a positive contribution to the energy transition. At the same time, this study offers practical implications for policymakers, identifying incentive system as a key driver of REC development, and for REC managers, highlighting the role of advanced benefit distribution systems in preserving community cohesion.

Keywords: Renewable Energy Communities; sustainability; economic benefits; systematic literature review; project evaluation

Paper type: Academic Research Paper

1 Introduction

Renewable Energy Communities (RECs) are a key instrument in the European regulatory framework for combating climate change and advancing the energy transition (Basilico et al., 2025; Biancardi et al., 2026; Moretti & Stamponi, 2023). Formally defined under the Renewable Energy Directive (RED II, 2018/2001/EU), RECs are legal entities characterised by the voluntary and open participation of citizens, small and medium-sized enterprises (SMEs), and local authorities in the collective self-consumption of electricity from renewable sources. RECs embody a predominantly mutualistic nature, in which financial profit is explicitly secondary to the generation of environmental, economic, and social benefits for their members and the local areas in which they operate, reflecting the principles of stakeholder theory (Freeman, 1984).

Although profitability is not the primary objective of RECs, financial sustainability remains a prerequisite for these entities to fulfil their broader mission (Biancardi et al., 2026; Martiniello & Presciutti, 2024). Without adequate economic returns, members lack the incentive to participate, and communities struggle to attract the investment needed to develop and maintain renewable energy infrastructure (Chaudhry et al., 2022b; Imset et al., 2025). The literature on RECs is rich and multidisciplinary, encompassing legal, social, engineering, and economic perspectives. Adopting a corporate finance lens, this paper focuses specifically on the economic and financial dimension of RECs, a challenging field that remains fragmented and underexplored.

To bridge this gap, this systematic literature review addresses the following research questions (RQs):

- ***RQ1:*** *What quantitative approaches have been used to assess the financial profitability of REC investments?*
- ***RQ2****: Do REC investments generate economic value?*

- ***RQ3:*** *What distribution mechanisms are employed to allocate economic benefits among REC participants?*

2 Methodology

This study adopts a systematic literature review to organise existing contributions on the economic dimension of RECs into a coherent framework, addressing the RQs through a replicable process of mapping and synthesising the available evidence (Tranfield et al., 2003). In particular, the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA 2020) guidelines (Page et al., 2021) were adopted. This model has gained widespread adoption in finance studies (Galletta et al., 2024), as it guarantees transparency and traceability across all phases of the review process: identification of records, screening of titles and abstracts, assessment of eligibility, and final inclusion of the sample.

In line with the PRISMA 2020 protocol, eligibility criteria were defined to ensure the relevance and quality of the review. The search strategy considered only original peer-reviewed articles written in English and published from 2018 onwards, the year of adoption of RED II (11 December 2018), while conference proceedings, book chapters, and grey literature were excluded. Articles were retrieved from Scopus, which was selected as the primary database following a preliminary search that revealed substantial overlap with Web of Science (WoS) results. The search query was structured in two steps based on keywords and Boolean operators.

The first step limits the studies to RECs. In the legal framework of the European Union, the term "energy communities" is used as an umbrella category that includes RECs, but also different configurations that are not within the scope of this research. For this reason, the search strategy includes both keywords "energy communit*" and "renewable energy communit*". A manual check was carried out during the screening phase to exclude articles referring to energy communities that do not present the characteristics required by RED II.

The second step included studies addressing the economic dimension of RECs, captured through keywords related to project evaluation metrics and profitability, such as "net present value", "internal rate of return", "discounted cash flow", "project financ*", "valuation methodolog*", "cash flow*", "payback period", "Public Private Partnership", "cost of equity", "scenario analys*", "sensitivity analys*", "discount rate", "cost of capital", "cost of debt", "capital structure", "leverage", and "corporate governance". Additional keywords related to economic benefits and their distribution among participants were also included, such as "financial benefit*", "economic benefit*", "revenue model*", "benefit sharing", and "benefit allocation". Studies addressing RECs exclusively from a technical, engineering, environmental, or social perspective, with no economic or financial dimension, were excluded.

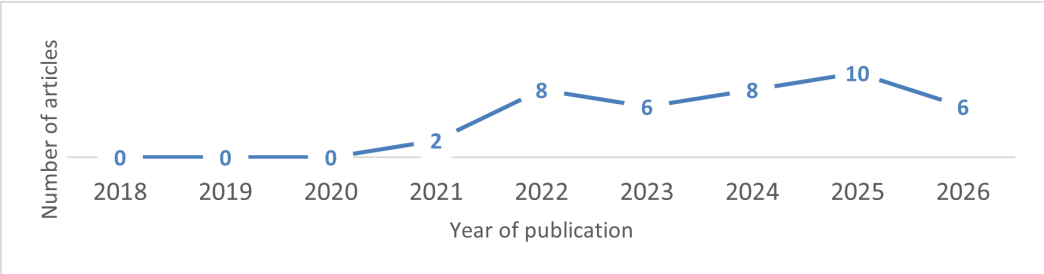


Fig. 1 Publication Trend.

The database search, finalised on 8 April 2026 following preliminary explorations, yielded 522 records. Upon the application of exclusion criteria based on peer review, language, and publication year, 292 articles underwent title and abstract screening, of which 53 proceeded to full-text review. Of these, 13 were excluded as their economic analysis proved marginal or absent. The final sample therefore consists of 40 articles, and Figure 1 illustrates the publication trend over the years. All phases of the review process, conducted in accordance with the PRISMA 2020 guidelines, are summarised in Figure 2.

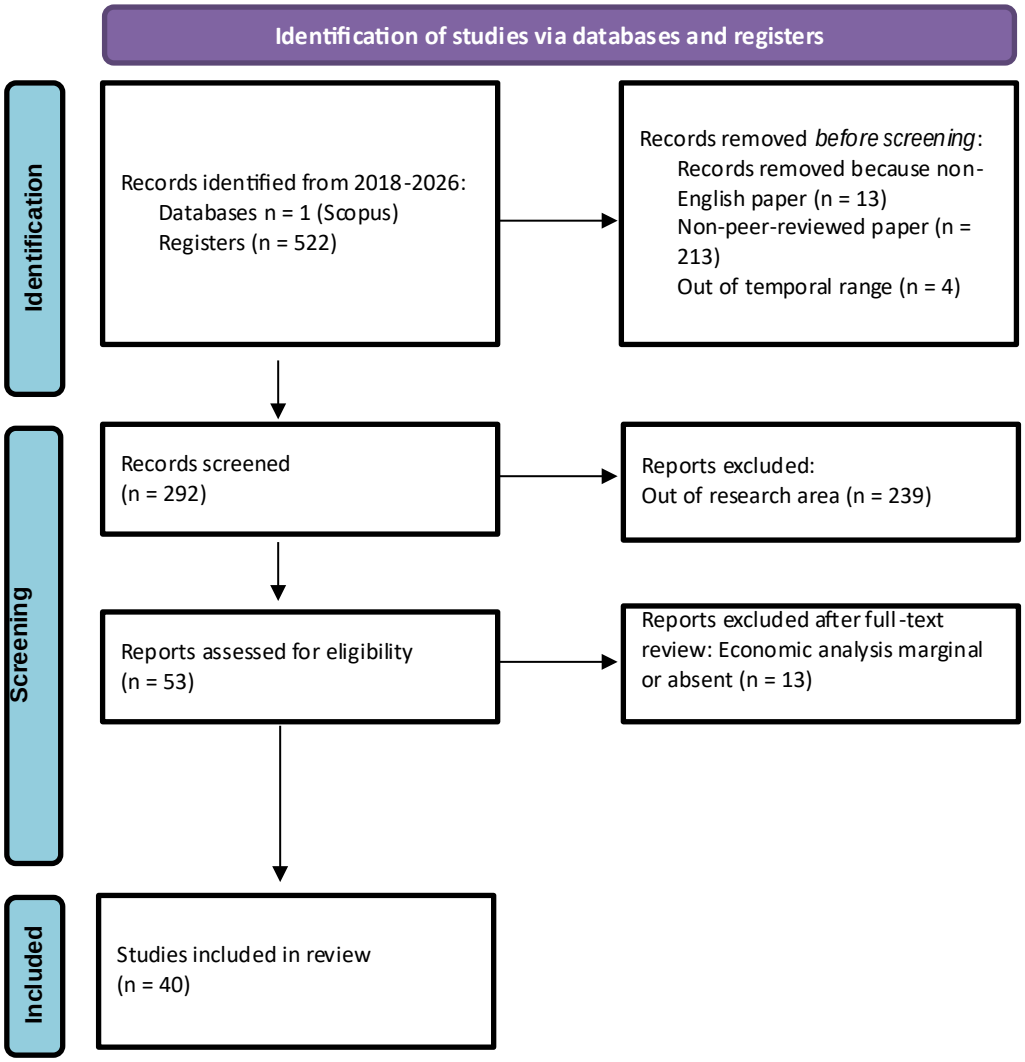


Fig. 2. PRISMA flow diagram for systematic literature reviews.

3 Results

3.1 Quantitative assessment of RECs economic performance

Regarding RQ1, most of the studies, about 60% included in this review, employ at least one capital budgeting technique, mainly based on on Net Present Value (NPV), Internal Rate of Return (IRR), and Payback Period (PBP). The empirical evidence addressing RQ2 is presented in the following sections.

3.1.1 Residential RECs

Among the various REC configurations studied in the literature, residential photovoltaic communities represent the most extensively investigated context, with private citizens as the primary participants. A significant share of this evidence draws on the Italian context, which has attracted considerable academic and investment interest owing to its advanced incentive system (Magni et al., 2025), consistently documenting positive economic returns.

Basilico et al. (2025) find that RECs generate positive NPV regardless of electricity price levels, confirming the robustness of returns across different market conditions. D'Adamo et al. (2023; 2024) corroborate these results, showing that NPV remains positive even under reduced incentive scenarios. Biancardi et al. (2026) show that economic sustainability holds even in the complete absence of public incentives, with positive NPV across all incentive scenarios. Sassone et al. (2025) document that profitability can be further enhanced by finding the optimal balance between prosumers and consumers, reducing the PBP.

Beyond the Italian context, the evidence on residential REC profitability is more mixed. Riquelme-Dominguez et al. (2025) find that forming a REC from scratch in Spain yields significantly higher collective savings and

increased NPV compared to individual self-consumption setups, driven by lower capital expenditure (CAPEX) and reduced maintenance costs. Meanwhile, Mutule et al. (2025) show that in Latvian RECs NPV can be negative due to high sensitivity to electricity market prices, concluding that without targeted incentives individual households face limited financial motivation to join a REC. In line with this evidence, Chaudhry et al. (2022a; 2022b) underline that over an extended investment horizon, a potential REC in Germany generates highly negative NPV values, driven by restrictive legislation and a structural revenue shortfall. This finding highlights the dependence of REC sustainability on public subsidies, a pattern further documented across multiple European regulatory contexts (Gjorgievski et al., 2023).

Overall, residential photovoltaic RECs prove to be profitable investments, though their financial viability remains dependent on public support. This dependence is more pronounced where regulatory frameworks are less developed, often requiring subsidies to remain financially attractive, while more advanced incentive systems such as the Italian one, which tailors support to plant size, geographical location, and market conditions, foster greater financial independence in the long term.

3.1.2 RECs with public and private involvement

Beyond the residential context, the literature observes that public and private involvement represent key drivers of value creation. Regarding mixed municipal and residential RECs in Italy, Ruggieri et al. (2023) report positive NPV over a 20-year horizon. Similarly, Moretti & Stamponi (2023) find significant positive NPV for a REC led by the municipality of Assisi, confirming that the inclusion of energy-intensive municipal facilities significantly amplifies economic returns. However, Cielo et al. (2021) document that mixed RECs present a positive investment potential that proves highly dependent on incentives. The authors show that in the presence of incentives the IRR exceeds the cost of capital, while in their absence it falls below it, rendering the investment financially unattractive.

Mari et al. (2025) extend this logic, highlighting that integrating commercial buildings into residential communities enhances NPV by exploiting the alignment between daytime commercial demand and solar energy production. This complementarity effect proves even more pronounced when REC composition is actively optimised. Blasuttigh et al. (2025) find that aggregating industrial firms based on load profile complementarity significantly increases NPV and reduces PBP, while Buonomano et al. (2024; 2026) confirm this approach in port and mixed urban configurations, achieving short PBP. Consistent results are documented in the Portuguese context, where Rego et al. (2025) find positive NPV for a mixed residential and private REC over a 25-year horizon.

Dujany et al. (2026) examine how governance configurations influence REC development in Italy, distinguishing between publicly and privately led initiatives. Their findings indicate that publicly led RECs pursue more inclusive and long-term oriented strategies, though their effectiveness is frequently constrained by bureaucratic complexity. Privately led RECs, by contrast, exhibit greater managerial flexibility and operational efficiency, but tend to prioritise short-term financial returns over the broader social objectives of the community. The authors consequently advocate for hybrid governance arrangements as these are best positioned to reconcile the social orientation of public leadership with the managerial agility of private actors. In line with this perspective, Martiniello & Presciutti (2024) examine Public-Private Partnership (PPP) models for RECs in Italy comparing two distinct configurations. The Cold model, where the public administration compensates the private concessionaire through an availability fee, generates a positive NPV. The Hot model, by contrast, struggles to reach financial equilibrium, as the private concessionaire assumes demand risk and requires substantial public capital contributions to remain viable. In both cases, once investment and management costs are absorbed, the residual economic benefits for individual community members remain minimal.

Taken together, these findings suggest that public and private participation consistently enhances REC economic performance through the complementarity of energy demand profiles. Furthermore, the governance model adopted plays a decisive role in determining REC development outcomes. However, the financial sustainability of these configurations remains dependent on public support.

3.1.3 RECs with mixed technology configurations

The literature also investigates RECs incorporating technologies beyond photovoltaics, showing that profitability is not uniformly achievable across all technological configurations. Ceglia et al. (2022) analyse a REC combining photovoltaic generation, biomass cogeneration, and hydroelectric power, finding that this integration reduces the PBP when favourable heat sale tariffs are in place. González et al. (2023) find that RECs based on local biomass district heating networks generate positive NPV only in the long term for urban configurations, contingent on a minimum community scale and high competing natural gas prices. In contrast to thermal

technologies, Marino et al. (2024) observe that the integration of battery storage slightly reduces NPV due to CAPEX and replacement costs, while substantially reducing energy withdrawals from the grid and thereby mitigating exposure to electricity price volatility. Consistent with these findings, Battaglia et al. (2024) find that electric vehicle recharging services systematically outperform standard REC. Meanwhile, Imset et al. (2025) underline that combining photovoltaics with battery storage and electric vehicles generates significant savings, though solar investments yield negative NPVs without balancing market revenues. Sousa et al. (2023) find that wind enhances REC profitability through temporal complementarity with photovoltaic. Moving to emerging technologies, Raimondi & Spazzafumo (2024) find that a hydrogen-based REC business model is not currently economically feasible, projecting sustainability only for large-scale configurations involving more than 3,000 people.

The literature indicates that alternative technology configurations hold potential for enhancing REC performance but remain financially constrained by high CAPEX. Further technological development is required before these solutions can be effectively integrated into REC business models, and future research should explore the conditions under which these configurations become financially sustainable.

3.2 Benefit redistribution mechanisms among REC participants

This section turns to RQ3, examining how the literature addresses the distribution of economic benefits among REC participants. Prosumers own the renewable energy installations and bear the upfront investment costs, while consumers benefit from energy sharing without owning any plant. Given these different roles, the two categories may develop diverging interests, in a dynamic analogous to the agency conflict between majority and minority shareholders (Jensen & Meckling, 1976). A flat distribution of costs and benefits is not optimal, as consumers could achieve greater financial benefits when energy is allocated based on their real time energy needs (Fina et al., 2022, Volpato et al., 2022), while maximising aggregate community benefits risks marginalising consumers in favour of prosumers (Volpato et al., 2024). This conflict becomes acute when the community composition changes, as new members with incompatible profiles can compromise existing members' returns, increasing the risk of participant dropout (Sousa et al., 2025).

A first response to this risk consists of mechanisms designed to ensure that every participant receives a minimum economic return from community membership. Manuel de Villena et al. (2022) address this through a centralised optimisation model that guarantee every participant a minimum share of locally produced energy and a minimum estimable financial return. Beraldi & Sener (2026) pursue a similar objective by introducing a safety net that guarantees each participant a minimum reduction in electricity bills of at least 8.17%, improving REC cohesion.

Other studies shift the focus from minimum return guarantees to the fairness of benefit distribution. Caferra et al. (2024) identify economic savings as the primary driver for REC participation, while Basilico et al. (2026) find that energy self-consumption is perceived as the fairest distribution criterion. Casalicchio et al. (2022a; 2022b) operationalise this notion through an optimal allocation method based on each participant's marginal contribution and a Fairness Index measuring distributional accuracy, which evaluates how closely a chosen business model aligns with the optimal allocation.

The literature also explores approaches to benefit distribution based on technology. Giudice et al. (2026) introduce a data driven mechanism that dynamically adjusts benefit allocation, rewarding members who align their behaviour with the community's energy objectives. Finally, Chantrel et al. (2021) develop a blockchain-based governance model for German RECs, enabling transparent and automated benefit redistribution among participants.

The distribution of economic benefits among REC members represents a governance challenge that communities must explicitly address, as the divergence of interests between prosumers and consumers can lead to participant dropout, undermining community stability. Flat allocation rules have proven insufficient to manage this complexity, and the literature has responded with mechanisms guaranteeing minimum returns, measuring distributional fairness, and leveraging technology for transparent benefit redistribution. However, identifying the optimal governance model requires further research, as the effectiveness of these mechanisms is likely to depend on broader contextual factors such as the regulatory and incentive framework in which RECs operate.

4 Discussion

This systematic review of the literature on RECs economic performance allows us to draw the following considerations.

First, given the multidisciplinary nature of RECs, disentangling the economic dimension from the technical one represents a methodological challenge. Nevertheless, the evidence reveals a clear dominance of NPV as the primary indicator for assessing REC profitability, complemented by IRR and PBP. Future studies should adopt a cost-benefit analysis as the primary evaluation tool, better capturing the environmental and social dimensions required under RED II.

Second, the empirical evidence reviewed suggests that RECs generally create value for private citizens, public administrations, and firms. However, the magnitude of returns varies considerably depending on several factors. The regulatory context emerges as one of the most impactful drivers: where incentive frameworks are insufficiently developed, profitability tends to decline significantly over time, while in supportive contexts such as Italy, RECs can remain viable even without public incentives. As the existing literature remains predominantly focused on single country analyses, future research should address this gap by conducting comparative studies across multiple countries with different regulatory frameworks.

Third, residential photovoltaic RECs represent the most extensively studied configuration in the literature, consistently documenting positive economic returns. The evidence further suggests that public and private involvement can significantly amplify these returns, leveraging the complementarity of energy demand profiles. Meanwhile, researchers should investigate the unexplored intersection between project finance and REC development.

Fourth, photovoltaic technology dominates REC configurations, though combining it with alternative technologies can potentially enhance community performance, despite these configurations not yet being fully profitable due to high CAPEX. Future studies should expand the literature on how public support can accelerate their deployment while fostering long-term financial independence.

Finally, the distribution of economic benefits among REC participants represents a governance challenge. The divergence of interests between prosumers and consumers can generate tensions leading to participant dropout. Flat allocation rules have proven insufficient, and for this reason further empirical research should explore how different distribution mechanisms perform over time, to identify the governance models most effective in preventing community attrition.

5 Conclusion and Implications

This paper presents a systematic review of the economic performance of RECs, based on a final sample of 40 articles. The evidence confirms that RECs generally represent financially viable investments, though their profitability is influenced by the regulatory framework and the technology adopted. Capital budgeting techniques dominate the literature, while the allocation of benefits between prosumers and consumers emerges as a critical governance challenge. These findings point to several implications for REC stakeholders.

From a practical standpoint, policymakers should design stable incentive schemes that gradually foster REC financial independence, as the regulatory context has proven a critical determinant of financial viability. Incentive frameworks should also be extended to cover alternative renewable technologies, supporting private citizens and firms in adopting energy solutions beyond photovoltaics. Finally, policymakers should consider developing a standardised evaluation methodology to make it immediately clear whether investing in a REC is worthwhile for potential participants.

From a managerial perspective, REC managers should invest in advanced benefit distribution systems and actively monitor member satisfaction, as flat benefit allocation rules have proven insufficient to prevent participant dropout and preserve community cohesion. PPP models represent a promising instrument for REC development, as hybrid arrangements have proven more effective than purely public or private configurations. Finally, private citizens, firms, and public administrations should consider joining RECs as economically sustainable investment vehicles, with mixed configurations proving particularly effective in generating significant economic value for all stakeholders.

Despite the potential limitations of this study, RECs are effectively contributing to the energy transition while generating value for their stakeholders.

6 Ethics Declaration

The authors state that no ethical clearance was required for the research presented in this paper. The authors also declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper

7 AI Declaration

The authors state that artificial intelligence tools were used for the translation and linguistic revision of English texts.

References

Basilico, P., Cartanese, G.M., D'Adamo, I., Ferrazza, M., Gastaldi, M. and Uricchio, A.F. (2026) "Renewable energy communities and energy poverty: Social and policy insights from empirical research in Italy", *Energy Policy*, Vol. 210, January, 115078, https://doi.org/10.1016/j.enpol.2026.115078.

Basilico, P., Biancardi, A., D'Adamo, I., Gastaldi, M. and Yigitcanlar, T. (2025) "Renewable energy communities for sustainable cities: Economic insights into subsidies, market dynamics and benefits distribution", Applied Energy, Vol. 389, March, 125752, https://doi.org/10.1016/j.apenergy.2025.125752.

Battaglia, V., Vanoli, L. and Zagni, M. (2024) "Economic benefits of Renewable energy communities in smart districts: A comparative analysis of incentive schemes for NZEBs", *Energy & Buildings*, Vol. 305, January, 113911, https://doi.org/10.1016/j.enbuild.2024.113911.

Beraldi, P. and Sener, N. (2026) "An integrated stochastic programming model for the fair and efficient operation of renewable energy communities", *Computers & Operations Research*, Vol. 191, March, 107442, https://doi.org/10.1016/j.cor.2026.107442.

Biancardi, A., D'Adamo, I., Donadel, A., Gastaldi, M. and Tavana, M. (2026) "Fostering sustainable economic development and mitigating energy poverty through renewable energy communities", *Energy, Sustainability and Society*, https://doi.org/10.1186/s13705-026-00564-0.

Blasuttigh, N., Negri, S. and Massi Pavan, A. (2025) "Optimal ATECO-Based Clustering and Photovoltaic System Sizing for Industrial Users in Renewable Energy Communities", *Energies*, Vol. 18, February, 763, https://doi.org/10.3390/en18040763.

Buonomano, A., Forzano, C., Giuzio, G.F., Maka, R., Palombo, A. and Russo, G. (2026) "Optimising renewable energy community aggregation for urban districts decarbonisation", *Renewable and Sustainable Energy Reviews*, Vol. 226, November, 116411, https://doi.org/10.1016/j.rser.2025.116411.

Buonomano, A., Giuzio, G.F., Maka, R., Palombo, A. and Russo, G. (2024) "Empowering sea ports with renewable energy under the enabling framework of the energy communities", *Energy Conversion and Management*, Vol. 314, July, 118693, https://doi.org/10.1016/j.enconman.2024.118693.

Caferra, R., Colasante, A., D'Adamo, I., Yilan, G. and Lancialonga, D. (2024) "A strategic analysis of renewable energy communities in achieving sustainable development", *Utilities Policy*, Vol. 90, August, 101810, https://doi.org/10.1016/j.jup.2024.101810.

Casalicchio, V., Manzolini, G., Prina, M.G. and Moser, D. (2022a) "Optimal Allocation Method for a Fair Distribution of the Benefits in an Energy Community", *Solar RRL*, Vol. 6, 2100473, https://doi.org/10.1002/solr.202100473.

Casalicchio, V., Manzolini, G., Prina, M.G. and Moser, D. (2022b) "From investment optimization to fair benefit distribution in renewable energy community modelling", *Applied Energy*, Vol. 310, February, 118447, https://doi.org/10.1016/j.apenergy.2021.118447.

Ceglia, F., Marrasso, E., Roselli, C., Sasso, M., Coletta, G. and Pellegrino, L. (2022) "Biomass-Based Renewable Energy Community: Economic Analysis of a Real Case Study", *Energies*, Vol. 15, August, 5655, https://doi.org/10.3390/en15155655.

Chantrel, S.P.M., Surmann, A., Erge, T. and Thomsen, J. (2021) "Participative Renewable Energy Community—How Blockchain-Based Governance Enables a German Interpretation of RED II", *Electricity*, Vol. 2, October, pp 471-486, https://doi.org/10.3390/electricity2040028.

Chaudhry, S., Surmann, A., Kühnbach, M. and Pierie, F. (2022a) "Renewable Energy Communities as Modes of Collective Prosumership: A Multi-Disciplinary Assessment, Part I—Methodology", *Energies*, Vol. 15, November, 8902, https://doi.org/10.3390/en15238902.

Chaudhry, S., Surmann, A., Kühnbach, M. and Pierie, F. (2022b) "Renewable Energy Communities as Modes of Collective Prosumership: A Multi-Disciplinary Assessment Part II—Case Study", *Energies*, Vol. 15, November, 8936, https://doi.org/10.3390/en15238936.

Cielo, A., Margiaria, P., Lazzeroni, P., Mariuzzo, I. and Repetto, M. (2021) "Renewable Energy Communities business models under the 2020 Italian regulation", *Journal of Cleaner Production*, Vol. 316, 128217, https://doi.org/10.1016/j.jclepro.2021.128217.

D'Adamo, I., Gastaldi, M., Koh, S.C.L. and Vigiano, A. (2024) "Lighting the future of sustainable cities with energy communities: An economic analysis for incentive policy", *Cities*, Vol. 147, January, 104828, https://doi.org/10.1016/j.cities.2024.104828.

D'Adamo, I., Mammetti, M., Ottaviani, D. and Ozturk, I. (2023) "Photovoltaic systems and sustainable communities: New social models for ecological transition. The impact of incentive policies in profitability analyses", *Renewable Energy*, Vol. 202, pp 1291-1304, https://doi.org/10.1016/j.renene.2022.11.127.

Manuel de Villena, M.M., Aittahar, S., Mathieu, S., Boukas, I., Vermeulen, E. and Ernst, D. (2022) "Financial Optimization of Renewable Energy Communities Through Optimal Allocation of Locally Generated Electricity", *IEEE Access*, Vol. 10, July, https://doi.org/10.1109/ACCESS.2022.3191804.

Dujany, N., Perotti, F.A. and Sohag, K. (2026) “Exploring governance for sustainable transitions in renewable energy communities: A comparative inquiry of private and public initiatives”, *Energy Policy*, Vol. 212, February, 115183, https://doi.org/10.1016/j.enpol.2026.115183.

European Union (2018) Directive (EU) 2018/2001 on the promotion of the use of energy from renewable sources, Official Journal of the European Union, L 328, pp 82–209. https://eur-lex.europa.eu/eli/dir/2018/2001/oj/eng

Fina, B., Monsberger, C. and Auer, H. (2022) “Simulation or estimation?—Two approaches to calculate financial benefits of energy communities”, *Journal of Cleaner Production*, Vol. 330, November, 129733, https://doi.org/10.1016/j.jclepro.2021.129733.

Freeman, R.E. (1984) *Strategic Management: A Stakeholder Approach*, Pitman, Boston.

Galletta, S., Mazzù, S., Naciti, V. and Paltrinieri, A. (2024) “A PRISMA systematic review of greenwashing in the banking industry: A call for action”, *Research in International Business and Finance*, Vol. 69, 102262, https://doi.org/10.1016/j.ribaf.2024.102262.

Giudice, R., Piscitelli, M.S. and Capozzoli, A. (2026) “A performance-based incentive sharing mechanism for communities of residential end users leveraging an ontology-driven approach”, *Building and Environment*, Vol. 287, October, 113786, https://doi.org/10.1016/j.buildenv.2025.113786.

Gjorgievski, V.Z., Velkovski, B., Minuto, F.D., Cundeva, S. and Markovska, N. (2023) “Energy sharing in European renewable energy communities: Impact of regulated charges”, *Energy*, Vol. 281, July, 128333, https://doi.org/10.1016/j.energy.2023.128333.

González, A., Arranz-Piera, P., Olives, B., Ivancic, A., Pagà, C. and Cortina, M. (2023) “Thermal energy community-based multi-dimensional business model framework and critical success factors investigation in the mediterranean region of the EU”, *Technology in Society*, Vol. 75, July ,102328, https://doi.org/10.1016/j.techsoc.2023.102328.

Imset, O.H., Ishmam, R., Kleinau, J., Lavrutich, M. and Yen, D.-B. (2025) “Financial viability of energy communities: The role of flexible assets, balancing market participation and peer-to-peer trading”, *Energy Economics*, Vol. 152, October, 108985, https://doi.org/10.1016/j.eneco.2025.108985.

Jensen, M.C. and Meckling, W.H. (1976) “Theory of the firm: Managerial behavior, agency costs and ownership structure”, *Journal of Financial Economics*, Vol. 3, October, pp 305-360, https://doi.org/10.1016/0304-405X(76)90026-X.

Magni, G.U., Bricca, D. and Familiari, S. (2025) “Economic Incentives for Renewable Energy Communities: a Scenario Analysis in the Transition Process Between the Experimental and Definitive Italian Policy Framework”, *Journal of Sustainable Development of Energy, Water and Environment Systems*, Vol. 13, No. 2, 1130589, https://doi.org/10.13044/j.sdewes.d13.0589.

Mari, C., Chierici, M., Franzò, S. and Causone, F. (2025) “Economic Performance Analysis of Jointly Acting Renewable Self-Consumption: A Case Study on 109 Condominiums in an Italian Urban Area”, *Buildings*, Vol. 15, August, 2808, https://doi.org/10.3390/buildings15162808.

Marino, C., Nucara, A., Panzera, M.F. and Pietrafesa, M. (2024) “Energetic and economic advantages of energy communities: a case study referred to a southern Italian town”, *Energy Reports*, Vol. 12, December, pp 6072-6081. https://doi.org/10.1016/j.egyr.2024.11.078

Martiniello, L. and Presciutti, A. (2024) “Implementation Models of RECs in Public-Private Partnerships: The Distribution of Risks and Benefits among the Participants in the Operation”, *Sustainability*, Vol. 16, August, 7358, https://doi.org/10.3390/su16177358.

Moretti, E. and Stamponi, E. (2023) “The Renewable Energy Communities in Italy and the Role of Public Administrations: The Experience of the Municipality of Assisi between Challenges and Opportunities”, *Sustainability*, Vol. 15, August, 11869, https://doi.org/10.3390/su151511869.

Mutule, A., Borscevskis, O., Astapov, V., Antoskova, I., Carroll, P. and Kairisa, E. (2025) “PV Energy Communities in Residential Apartments: Technical Capacities and Economic Viability”, *Sustainability*, Vol. 17, March, 2901, https://doi.org/10.3390/su17072901.

Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., Mulrow, C.D., Shamseer, L., Tetzlaff, J.M., Akl, E.A., Brennan, S.E., Chou, R., Glanville, J., Grimshaw, J.M., Hróbjartsson, A., Lalu, M.M., Li, T., Loder, E.W., Mayo-Wilson, E., McDonald, S., McGuinness, L.A., Stewart, L.A., Thomas, J., Tricco, A.C., Welch, V.A., Whiting, P. and Moher, D. (2021) "The PRISMA 2020 statement: an updated guideline for reporting systematic reviews", BMJ, Vol. 372, March, n71, https://doi.org/10.1136/bmj.n71.

Raimondi, G. and Spazzafumo, G. (2024) "Integrating renewable energy communities and Italian UVAM project through renewable hydrogen chain", e-Prime - Advances in Electrical Engineering, Electronics and Energy, Vol. 10, December, 100819, https://doi.org/10.1016/j.prime.2024.100819.

Rego, N., Castro, R. and Lagarto, J. (2025) “Sustainable energy trading and fair benefit allocation in renewable energy communities: A simulation model for Portugal”, *Utilities Policy*, Vol. 96, June, 101986, https://doi.org/10.1016/j.jup.2025.101986.

Riquelme-Dominguez, J.M., Sempértegui, M.E., Roldan-Fernandez, J.M., Serrano-Gonzalez, J. and Riquelme-Santos, J.M. (2025) “Economic assessment of self-consumption and energy communities: Profit distribution insights from a real case study”, *Energy Conversion and Management: X*, Vol. 28, August, 101198, https://doi.org/10.1016/j.ecmx.2025.101198.

Ruggieri, G., Gambassi, R., Zangheri, P., Caldera, M. and Verde, S.F. (2023) “Key Economic Drivers Enabling Municipal Renewable Energy Communities' Benefits in the Italian Context”, *Buildings*, Vol. 13, November, 2940, https://doi.org/10.3390/buildings13122940.

Sassone, A., Ahmed, S., Ciocia, A., Malgaroli, G. and D'Angola, A. (2025) "The Role of Participant Distribution and Consumption Habits in the Optimization of PV Based Renewable Energy Communities", *Energy Engineering*, Vol. 122, No. 5, April, pp 1715-1733, https://doi.org/10.32604/ec.2025.058781.

Sousa, J., Lagarto, J., Camus, C., Viveiros, C., Barata, F., Silva, P., Alegria, R. and Paraíba, O. (2023) "Renewable energy communities optimal design supported by an optimization model for investment in PV/wind capacity and renewable electricity sharing", *Energy*, Vol. 283, July, 128464, https://doi.org/10.1016/j.energy.2023.128464.

Sousa, J., Perinhas, S., Viveiros, C. and Barata, F. (2025) "Increasing Economic Benefits in Renewable Energy Communities with Solar PV and Battery Storage Technologies: Insights from New Member Integration", *Energies*, Vol. 18, September, 4815, https://doi.org/10.3390/en18184815.

Tranfield, D., Denyer, D. and Smart, P. (2003) "Towards a methodology for developing evidence-informed management knowledge by means of systematic review", *British Journal of Management*, Vol. 14, September, pp. 207–222, https://doi.org/10.1111/1467-8551.00375.

Volpato, G., Carraro, G., Cont, M., Danieli, P., Rech, S. and Lazzaretto, A. (2022) "General guidelines for the optimal economic aggregation of prosumers in energy communities", *Energy*, Vol. 258, July, 124800, https://doi.org/10.1016/j.energy.2022.124800.

Volpato, G., Carraro, G., Dal Cin, E. and Rech, S. (2024) "On the Different Fair Allocations of Economic Benefits for Energy Communities", *Energies*, Vol. 17, September, 4788, https://doi.org/10.3390/en17194788.

Rethinking Product Data Management Through Digital Product Passports: Data Governance for Global Supply Chains

Martina De Giovanni, Antonella Marsano, Mariangela Lazoi
University of Salento, Lecce, Italy

Abstract

Increasing regulatory pressures and societal expectations have made digital transparency a critical requirement, positioning DPPs as key instruments to support traceability and supply chain-based decision-making. However, their effectiveness depends not only on technological infrastructure (e.g., blockchain, RFID, IoT sensors, and system integration), but also on robust data governance frameworks.

This study aims to explore the role of data in enabling the effective implementation of Digital Product Passports (DPPs) as strategic governance tools for transparency, accountability, and sustainable value creation in global supply chains. The research is based on a Systematic Literature Review (SLR) conducted through Scopus and Web of Science, focusing on the evolution of Product Data Management (PDM) and its transition toward the emergence of DPPs. While early PDM approaches emphasized manufacturer-controlled data for operational efficiency, more recent developments show a shift toward multi-actor data ecosystems incorporating environmental, social, and end-of-life information. This transition has been accelerated by European policy initiatives, particularly the Circular Economy Action Plan (CEAP) and the Ecodesign for Sustainable Products Regulation (ESPR), which are transforming DPPs from voluntary tools into mandatory regulatory instruments.

DPPs are especially relevant in high-impact sectors (e.g., electronics, textiles, construction), where they contribute to reducing environmental footprints and enhancing supply chain accountability.

The literature review identifies forty-two data typologies (referred to as factors) required for DPP implementation and organizes them into ten key data categories. The proposed data mapping framework is linked to the Environmental, Social, and Governance (ESG) criteria, emphasizing the integration of data into governance processes to support strategic decision-making. DPPs enable continuous monitoring and lifecycle traceability fostering alignment between circularity and sustainability practices. Moreover, they facilitate bidirectional information flows: consumers gain access to reliable product data, while their feedback can inform producers and regulators, promoting adaptive governance and collective learning.

Despite these opportunities, significant challenges remain, including data asymmetries, privacy concerns, lack of standardization, and unclear stakeholder responsibilities. Without well-defined governance structures, DPPs risk reinforcing existing power imbalances rather than enhancing transparency.

The study concludes that governance is a critical determinant of DPP success. It calls for the development of scalable, harmonized, and compliant data governance models capable of supporting DPP adoption across sectors while ensuring data continuity along the value chain. By framing governance as a strategic enabler, DPPs can evolve beyond compliance tools into foundational mechanisms for circular and sustainable value creation.

Keywords: Corporate Governance, Product Data Management, Digital Transparency, Sustainable Supply Chains

Paper type: Academic Research Paper

1 Introduction

The transition to sustainable and circular economic models is redefining data's role in global supply chains. Growing regulatory pressure, particularly from European initiatives like the Green Deal and the Circular Economy Action Plan (CEAP), requires companies to account for environmental and social impacts throughout the entire life cycle (Mollahassani et al., 2025; Papile et Del Curto, 2025).

In this landscape, the Digital Product Passports (DPPs) emerge as a key tools for traceability and transparency integrating value chain for stakeholders including companies, regulators, consumers, and recyclers (Gieß et Möller, 2025; Jeganathan et Szymkowiak, 2025). This reflects Product Data Management (PDM)'s evolution from company-centric systems to multi-stakeholder ecosystems supporting sustainability (Fares et al., 2025). DPP development followed a conceptual phase (2020–2022) followed by an implementation phase (from 2022 onward).

Existing studies focus on enabling technologies, such as blockchain, the Internet of Things (IoT), and systems integration (Kebede et al., 2024), while neglect data aligned with environmental, social, and governance (ESG) objectives (Jensen et al., 2024). This gap is particularly critical because the DPP effectiveness depends on structuring data across environmental (eco-design, LCA), social (due diligence, human rights), and governance (transparency, risk management) dimensions (Gieß et Möller, 2025; Kim et al., 2023). This approach enables the

integration of sustainability metrics into strategic decision-making processes and promotes alignment between circularity and sustainability practices (Ospital et al., 2023; Pracucci et Giovanardi, 2025). In this context, the study aims to investigate how product data can be structured and governed within DPP systems to effectively support traceability, sustainability, and life-cycle-based analyses. To address this objective, a systematic literature review (SLR) is conducted to identify the key data categories involved in DPP implementation and to develop a comprehensive framework for their classification and governance across the supply chain. This paper contributes to the literature by (i) bridging the gap between DPP technological architectures and ESG-oriented data governance, (ii) providing a structured mapping of product data across life cycle stages, and (iii) outlining the governance implications for the effective implementation of DPPs in complex, multi-stakeholder contexts.

2 Research Methodology

The objective of the paper is to understand how to structure data governance to make the DPP an effective tool for transparency, traceability and sustainability across supply chains.

Therefore, a Systematic Literature Review (SLR) is used as the research methodology to clarify the scientific body of knowledge through a replicable, scientific and transparent protocol that aims to minimize bias through an exhaustive literature search (Tranfield et al., 2003). The systematic approach provides a comprehensive, unbiased synthesis of many relevant studies of the literature (Egger et al., 2001).

Protocols include aspects related to the design and implementation of the review, in line with the definition of research questions, criteria for documents eligibility, search strategy, study selection, data extraction, contents assessment, synthesis and reporting. A protocol provides a step-by-step guide to carrying out the work in a clear and consistent way. Several procedures for outlining an effective review protocol have been identified in the literature and, although they are different, they have some commonalities.

The methodological SLR protocol used in this paper is based on (Corallo et al., 2023), which consists of four steps: (i) review planning, (ii) review execution, (iii) analysis and (iv) reporting. Each phase will be discussed in more detail in the following subsections and the procedure was summarized in Figure 1.

2.1 Review planning

In the initial phase of the research, it is essential to define a clear objective to guide the entire analysis. Therefore, in line with the objective of this study, the systematic literature review (SLR) aims to identify all relevant types of supply chain data enable to support DPP in all industrial sectors involved, with a focus on the evolution of PDM and its transition toward the emergence of DPPs.

Consequently, the research question (RQ) formulated for this study is: What product data must be managed and integrated into digital systems, such as the Digital Product Passport in manufacturing industries to support traceability and sustainability across supply chains?

The Scopus and Web of Science (WoS) databases were selected as reference sources for scientific contributions, thanks to their reliability and broad interdisciplinary coverage.

To facilitate the implementation of the review, some appropriate keywords were selected to ensure a comprehensive and targeted literature search.

In addition, eligibility criteria were established for the inclusion or exclusion of articles identified in the literature. Four inclusion criteria were selected for this study as reported in Figure 1. However, no time restrictions were applied.

Finally, a preliminary qualitative content analysis was adopted as mechanism for article selection to align the content with the analysis scope.

2.2 Review execution

This phase is based on the implementation of the SLR. Scopus and WoS have been interrogated in November 2025 by selecting keywords appropriately combined through Boolean operators The search query obtained is shown in Figure 1. Papers can be selected only if keywords are included within the title, abstract or the keywords of the paper. Specifically, as research field has been selected TIT-ABS-KEY in Scopus is, while topics (TS) in WoS that refers to the same fields. This formula ensured that the review captured a full range of interest within literature.

Moreover, the initial sample was refined by applying specific filters in accordance with the inclusion and exclusion criteria as detailed in Figure 1 together with the quantitative results of initial sample.

2.3 Analysis

The third step aims to a comprehensive analysis of the papers returned from the search. The quantitative advancement of the articles analyzed during the SLR procedure is reported in Figure 1.

A systematic qualitative content analysis was carried out by organizing and synthesizing the information gathered from the selected documents in a structured predefined excel matrix. The relevant information extracted referred to bibliometric information, the study's objective, main findings, and finally core contents such as data types processed in each study, technologies supporting DPP implementation and PDM background.

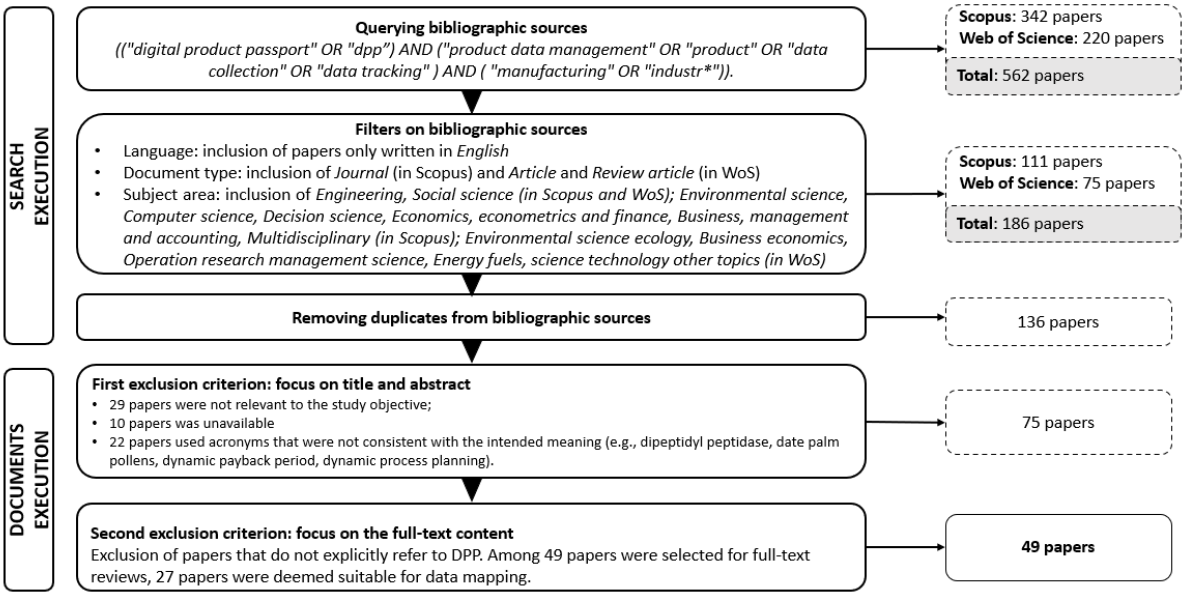


Figure 1. SLR procedure.

2.4 Reporting

This phase provides an overview about the most relevant findings emerged from the body of knowledge, highlighting key relationships and insights related to the research topic. The analysis content in Section 3 is mainly organized into two sections. The first one examines the evolution of PDM by identifying the key factors that have driven it with particular emphasis on enabling technologies. The second one develops a framework for mapping data to be included in the DPP, encompassing various sectors. Finally, such data was discussed about their implications from a governance perspective.

3 Content analysis

3.1 Evolution of product data management in digital product passport

Historically, PDM has been defined as a single repository of heterogeneous data related to the entire product lifecycle, with utility limited to the boundaries of the organization. In this context, data (such as CAD drawings, bills of materials, and analysis reports) served primarily to optimize the design and production phases (Do, 2017; Liu et al., 2021). However, this model led to the creation of information silos, where data remained trapped in fragmented ecosystems and a persistent reliance on paper-based documentation. All of this resulted in high interoperability costs between different systems, difficulties in traceability throughout the entire supply chain, and risks related to intellectual property (Mabkhot et al., 2025). To overcome fragmentation, research introduced the concepts of Digital Twin and Digital Thread. While the Digital Twin provides a synchronized virtual representation of the physical asset to monitor performance, the Digital Thread aims to link the data flow across all phases of the lifecycle. Despite this approach, however, these tools have often remained limited to ad hoc applications and focused on operational benefits for the manufacturer, with poor interoperability among different supply chain actors (Mabkhot et al., 2025). The introduction of the DPP, driven by European regulations such as the ESPR, marks the definitive evolution of the PDM into a tool for universal transparency (Langley et al., 2023). The DPP is no longer just an internal technical data sheet, but a 'digital identity card' that accompanies the product throughout its entire lifecycle (Gösken, 2025; Petrik et al., 2025). This evolution is driven by the need to bridge the information gaps that hinder the circular economy, making data actionable not only for businesses

but for all stakeholders (Kemmner et al., 2025; Mollahassani et al., 2025). The true transformation lies in expanding the user base; in fact, the DPP acts as a 'boundary object' that facilitates a common language among different actors (Gieß et Möller, 2025; Petrik et al., 2025). In particular, among these we can identify consumers, who, through the DPP, are able to make informed purchasing decisions based on verifiable data regarding environmental footprint, durability, and reparability (Gieß et Möller, 2025; Ruismäki et al., 2025); repairers and recyclers, who will have detailed instructions on disassembly and material composition—information essential for the recovery of raw materials (Carvalho et al., 2025); and authorities, for whom it enables verification of compliance with environmental and social regulations (Gösken, 2025; Langley et al., 2023) and for the supply chain, ensuring collaboration among suppliers and improving risk management and the traceability of raw materials (Abreu et al., 2025; Hakola et al., 2025).

Data management has thus evolved from a technical support activity into a strategic lever for social and environmental responsibility, where data becomes a digital common necessary to ensure global sustainability (Kemmner et al., 2025; Mollahassani et al., 2025). Therefore, among the key drivers of this evolution, we can certainly identify the European Union's legislative package related to the Green Deal and the CEAP. In particular, the ESPR regulation mandates the creation of the DPP as a mandatory requirement for placing products on the European market, starting with high-impact sectors such as electric batteries and textiles, requiring data that is not only collected but also accessible and verifiable (Jensen et al., 2024; Petrik et al., 2025; Saenz et al., 2024). In addition to regulatory requirements, the DPP was created to bridge the information gap across the supply chain by centralizing data from various stakeholders into a single digital identity, overcoming the fragmentation of the traditional model (Cruz et al., 2025; Kebede et al., 2024; Wan et Jiang, 2025). Furthermore, changing purchasing behaviour has required greater transparency regarding origin, working conditions, and carbon footprint to avoid greenwashing (King et al., 2023; Ospital et al., 2023; Papile et Del Curto, 2025; Ruismäki et al., 2025), as also demonstrated by empirical studies showing that consumers are willing to pay a premium of up to 17% for products bearing a label certifying their sustainability (Colasante et al., 2025; D'Adamo et al., 2025), serving as a bridge of trust between brands and the public (Alves et al., 2023; Carvalho et al., 2025; Jeganathan et Szymkowiak, 2025). As a further key aspect, it is impossible to implement circular economy strategies without precise technical data about product end-of-life (Mollahassani et al., 2025).

The literature also points to a shift from mere operational efficiency to a human-centric approach (Voulgaridis et al., 2024). The DPP has also been adopted to protect intellectual property (Gösken, 2025) and, through the use of blockchain, allows the product's history to be recorded in an immutable manner (Colasante et al., 2025; Jensen et al., 2023; Panza et al., 2023).

3.2 Technology architecture and digital pillars for data governance and transparency

The literature reviewed identifies an integrated technological architecture that serves as the cornerstone for the digital governance of the DPP. In particular, recent research highlights that the interaction between product data and advanced digital tools is essential for ensuring transparency, accountability, and sustainability in global supply chains (Abreu et al., 2025; De Diego et Gutiérrez-Aguero, 2025; Papile et Del Curto, 2025). Among the technologies playing a key role in data governance, blockchain stands out (Abreu et al., 2025). It is considered the technology of choice for establishing a 'digital chain of trust'; it provides a decentralized and immutable ledger where every transaction or product transformation is chronologically recorded, preventing tampering and fraud. Thanks to the use of Smart Contracts (Liu et al., 2021; Voulgaridis et al., 2024), it enables the automation of governance rules, verifying the identity of actors and managing access rights in a granular manner; meanwhile, integration with Self-Sovereign Identity (SSI) systems (De Diego et Gutiérrez-Aguero, 2025; Langley et al., 2023) ensures that companies retain control over their sensitive data, sharing only the information necessary for regulatory compliance. On the other hand, IoT enables the automated and continuous collection of data from the field through the use of smart sensors that monitor critical parameters—such as energy consumption—in real time, transforming the product passport from a static record into a dynamic document (De Diego et Gutiérrez-Aguero, 2025; Fares et al., 2025; Kemmner et al., 2025). Technologies such as RFID and NFC overcome the limitations of traditional barcodes, enabling the unique identification of individual garments or components and facilitating tracking throughout the entire lifecycle, from production to end-of-life (Hakola et al., 2025; Wan et Jiang, 2025). Furthermore, to avoid creating new silos, the DPP must be integrated with existing IT infrastructures, and the adoption of modular ontologies standardizes data formats, allowing heterogeneous systems (ERP, SCM) (Alves et al., 2023; Cruz et al., 2025; King et al., 2023) to communicate with one another without technical barriers. Integration with enterprise management systems reduces the administrative burden and the risk of human error resulting from manual data entry (Petrik et al., 2025).

4 Mapping Data Across the Supply Chain: A Proposed Framework for DPP

In this study, the analysis focused on identifying and organizing the data generated and exchanged throughout the supply chain, with particular attention to its potential contribution to the development of the DPP. This classification (shown in Figure 2) has made it possible to map the relevant data across the supply chain in a coherent and systematic manner, providing a solid foundation for the definition and implementation of the DPP. This analysis outlined the sectors affected by the introduction of the DPP are the following: constructions (Çetin et al., 2023; Kebede et al., 2024; Pracucci et Giovanardi, 2025; Ruismäki et al., 2025; Vergés et al., 2026), textile (Alves et al., 2023; Carvalho et al., 2025; Colasante et al., 2025; Cruz et al., 2025; D'Adamo et al., 2025; Domskienė et Gaidule, 2024; Gonçalves et al., 2025; Ospital et al., 2023; Papile et Del Curto, 2025; Won, Y.; Jang, S.; Kim, H.Y.; Lee, Y.원 et al., 2025), manufacturing (Jensen et al., 2023, 2024; Mollahassani et al., 2025; Panza et al., 2023), electronics (Saenz et al., 2024), energy (Boukhatmi et al., 2023), EV battery (Kim et al., 2023). Moreover, in some studies the resulting data are regarded as cross-sector since it serves multiple industries (De Diego et Gutiérrez-Aguero, 2025; Fares et al., 2025; Gieß et Möller, 2025; Kühn et al., 2025; Wan et Jiang, 2025).

A total of 42 information factors were identified and subsequently organized into 10 categories to facilitate the data interpretation and integration.

Moreover, such categories have been further linked to ESG criteria. The environmental (E) perspective encompasses five categories. The first category "Material & Composition" includes data regarding the nature of raw materials, including virgin and circular-origin materials, bio-based characteristics, as well as the product's composition with particular attention to the presence and quantity of chemicals and the bill of materials. On the other hand, End-of-Life (EoL) data which includes instructions for disassembly and disposal, the amount of waste generated, recovery strategies (R-strategies), and the product's residual economic value was included in the "EoL & Circularity" category. Then, "Environmental Performance" category involved indicators and impacts by considering emissions and liquid effluent volume according to a circular and non-renewable approach.

Some factors related to energy consumption during the manufacturing process as well as the energy consumption in service was considered respectively within "Production" category and "Use Phase & Performance" one. The rest of factors included in those categories were referred respectively to Governance (G) and Social (S) criteria.

Therefore, the Governance (G) perspective encompasses a total of five macro-categories. "Design & Technical" category collects information regarding the physical and technical properties of the product, including digital models (e.g., CAD, BIM). Another key area is "Product Identification", which encompasses all identifying information related to suppliers and manufacturers, including identity, ownership, product code, product category, and the location of the production site. On the other hand, the "Traceability" category focuses on the data related to packaging and information needed to ensure product traceability throughout its supply chain by considering both past-oriented information (tracing) as well as present & forward information (tracking). Finally, the "Deployment & Support" category includes product warranty factor, whereas the remaining factors support users and are therefore aligned with the Social (S) lens.

Finally, in addition to the abovementioned categories, the "Social Impact & Compliance" category was included from a Social (S) point of view too. It includes ethical and social responsibility aspects related to working conditions, hours worked, and corporate policies as well as certification standards.

CATEGORY	FACTORS \ REFERENCES	Vergès et al., 2025	Ruismäki et al., 2025	Colasante et al., 2025	D'Adamo et al., 2025	Papile et al., 2025	Ferreira Cruz et al.,	Pracucci et al., 2025	Gonçalves et al., 2025	Gieß et al., 2025	Carvalho et al., 2025	Wan et al., 2025	Kühn et al., 2025	Won et al., 2025	Fares et al., 2025	De Diego et al., 2025	Juresa et al., 2025	Jensen., 2024	Kebede et al., 2024	Saenz et al., 2024	Alves et al., 2024	Domskiene et al., 2024	Boukhatmi et al., 2023	Catin et al., 2023	Panza et al., 2023	Jensen et al., 2023	Kim et al., 2023	Ospital et al., 2023
Materials & Composition	Virgin Raw Material	X			X					X		X			X	X			X									X
	Recycled/reused material	X				X	X		X				X					X	X			X	X	X				
	Composition (%)						X		X	X	X	X	X	X			X	X	X			X	X	X	X	X		X
	Bio-based Characterization		X			X		X														X		X				
	Recovered/Recyclable quantity		X				X					X												X			X	
	(Non-) hazardous Chemical quantity					X	X	X	X		X	X			X		X	X	X				X	X	X		X	
	Bill of materials											X					X		X				X	X		X		
Design & Technical	CAD/ BIM model											X							X				X	X				
	Physical properties					X				X	X	X	X	X					X					X	X	X	X	
Production	Data from sensors					X								X					X									
	Production Date															X								X		X		
	Quality control		X					X				X		X					X		X		X	X		X		X
	Energy Consumption					X	X	X	X			X	X						X						X			
	Production Process			X			X	X	X	X		X	X	X			X	X								X	X	X
Product Identification	Supplier identity							X										X	X									X
	Ownership					X					X								X					X				
	Product ID					X					X		X			X	X	X	X				X	X				
	Product category											X				X							X	X				
	Manufacturer address					X	X				X	X	X	X	X	X						X				X		X
	Manufacturer identity													X		X							X	X		X		
Traceability	Tracking information					X		X	X	X		X		X				X	X					X		X		X
	Packaging									X		X							X							X		
	Tracing information										X		X	X					X					X			X	
Use Phase & Performance	Use Instructions/Safety warnings									X		X					X							X		X		
	User feedback													X			X	X										
	Expected lifespan											X	X											X		X		
	Energy consumption in service	X	X														X		X									
Deployment &Support	Product Warranty					X											X						X					
	Installation instructions																							X		X		
	After sale service					X				X	X	X	X				X	X	X				X	X				
EoL &Circularity	(Dis) Assembly instructions	X	X			X				X	X	X				X			X	X		X	X	X		X		
	Disposal instructions	X	X	X		X				X	X			X			X	X	X	X		X		X		X		
	Waste quantity						X												X								X	
	Initiative R strategies		X		X	X	X			X	X	X	X	X			X		X	X			X	X				
	EoL economic value					X						X								X				X	X		X	
Environmental Performance	Non-Renewable liquid effluent volume						X		X															X	X			
	Recycled liquid effluent volume						X																					
	Environmental impatcs	X	X					X	X		X	X	X	X			X	X	X		X	X		X			X	X
	Performance indicators	X					X	X											X					X		X		X
	Emissions	X		X		X	X	X		X		X	X	X			X							X	X			
Social Impact & Compliance	Labor Practices & Social Responsibility		X	X	X					X	X			X			X	X			X				X	X	X	X
	Certifications standards		X		X			X		X	X		X		X	X	X		X				X	X		X		X

Construction
Textile
Cross-sector
Manufacturing
Electronics
Energy
EV Batteries

Figure 2. Data Mapping Framework for DPP according to ESG criteria.

5 Discussion and Conclusions

Since the various stages of the value chain involve a multitude of stakeholders, each responsible for the production, management, or transmission of specific information, a holistic approach is necessary to capture the heterogeneity and distributed nature of these data flows. In this respect, the research highlights how the digitalization of information can translate regulatory requirements into concrete operational advantages, by structuring data according to the environmental, social, and governance (ESG) dimensions.

It is important to highlight that, when considering data mapping from a multi sector perspective, not all data categories hold the same level of relevance across industries. Some types of data emerge as broadly applicable, supporting several sectors with relatively standardized structures, while others are more tightly linked to specific industrial contexts and operational characteristics. This distinction reflects underlying structural and functional differences between industries, which influence both the relevance and the level of detail of product-related data throughout its supply chain. Specifically, among those data (factors) related to highly formalized digital design representations (e.g., CAD/BIM models) are predominantly relevant in sectors like construction or advanced manufacturing, where design processes rely heavily on standardized digital modeling. Similarly, data such as detailed bill of materials, production date, expected lifespan, operating and safety instructions, and installation or use instructions tend to be more prominent in sectors characterized by high levels of standardization, regulatory requirements, and product complexity. Their reduced presence in other sectors may reflect differences in production processes, documentation practices, and perceptions of product durability and use conditions. Among data concerning "Traceability" category, the highest relevance is addressed to packaging which can also vary significantly depending on the sector, influenced by factors such as logistics complexity, regulatory constraints, and its role along product value chains. Likewise, in the "Use phase & Performance" category, factors such as product installation and energy consumption in service are critical in some industries (e.g., energy-using products or technical equipment), while being marginal or not directly applicable in others. Moreover, from the analysis is emerged a clear emphasis on some factors that represent a converging baseline of product information requirements driven by sustainability regulations, circular economy objectives, and the need for greater transparency and traceability. Specifically, they include environmental impact data, certification standards, and information supporting circular economy R-strategies (e.g., reuse, repair, remanufacturing,

recycling). Strong attention is also given to disassembly and disposal instructions, product composition data, and the identification and quantification of hazardous (and non-hazardous) chemical substances.

Although the importance of sustainability aspects to be included within the DPP is growing, the existing literature highlights significant gaps in the lack of systematic and scalable methods to structure heterogeneous data for Environmental Product Declarations (Vergés et al., 2026). Furthermore, the effectiveness of circular models strongly depends on the adoption of eco-design principles, including repairability and the use of innovative materials such as bio-based fibers and recycled components. These aspects should be considered elements in an environmental impact assessment conducted through Life Cycle Assessment (LCA) analysis.

In addition, the social perspective of DPP must be considered to ensure social responsibility along the supply chain by including the previously mapped factors in line with the Corporate Sustainability Due Diligence Directive (CSDDD).

Data gaps do not merely represent sectoral limitations, but concrete opportunities for a systemic innovation. From a managerial perspective, companies should prioritize tailored approaches rather than one-size-fits-all models by integrating data on product documentation, use-phase impacts, and user interactions. By enhancing transparency and traceability, governance needs are addressed, ensuring verifiable information on product origin, material composition, and end-of-life management. Therefore, investing in digital infrastructures, data standardization, and enhanced supply chain collaborations will strengthen data management practices. The benefits are multiple: more accurate sustainability assessments, facilitated regulatory compliance, and support for informed decisions along the value chain. Ultimately, a strategic view of the DPP is essential: it not only provides structured information but empowers consumers for informed choices, recyclers for material recovery, regulators for standards enforcement, and generates new opportunities in traceability, resale, repair, and recycling.

It is also important to note that the analyzed literature reflects a relatively recent and rapidly evolving research field. Therefore, some identified gaps can be partially attributed to the early stage of DPP development as it is still evolving and will become progressively mandatory under upcoming regulatory requirements in the coming years.

Future research should investigate how sectoral specificities shape DPP data requirements, develop interoperable data models, and empirically validate the integration of missing data categories to measure their impact on sustainability performance.

6 Ethics declaration

The authors have to declare any ethical issues.

7 AI declaration

AI tool was employed as a graphical aid to generate visual elements, which were then manually selected and integrated into the data mapping framework proposed. All other textual or analytical contents of the paper were produced independently.

References

Abreu, Hugo ; Pereira, Vasco et Barata, João. (2025). Blockchain-based digital product passport: design principles and demonstration. *International Journal of Production Research*, 63(16), pp. 5863-5882.

Alves, Luís ; Sá, Miguel ; Cruz, Estrela Ferreira ; Alves, Toni ; Alves, Marcelo ; Oliveira, João ; Santos, Manuel et Rosado Da Cruz, António Miguel. (2023). A Traceability Platform for Monitoring Environmental and Social Sustainability in the Textile and Clothing Value Chain: Towards a Digital Passport for Textiles and Clothing. *Sustainability*, 16(1), pp. 82.

Boukhatmi, Ässia ; Nyffenegger, Roger et Grösser, Stefan N. (2023). Designing a digital platform to foster data-enhanced circular practices in the European solar industry. *Journal of Cleaner Production*, 418, pp. 137992.

Carvalho, Catarina ; Silva, Carla Joana et Abreu, Maria José. (2025). Circular Economy: Literature Review on the Implementation of the Digital Product Passport (DPP) in the Textile Industry. *Sustainability*, 17(5), pp. 1802.

Çetin, Sultan ; Raghu, Deepika ; Honic, Meliha ; Straub, Ad et Gruis, Vincent. (2023). Data requirements and availabilities for material passports: A digitally enabled framework for improving the circularity of existing buildings. *Sustainable Production and Consumption*, 40, pp. 422-437.

Colasante, Annarita ; D'Adamo, Idiano ; Desideri, Simona ; Iannilli, Martina et Mangani, Valeria. (2025). Environmental Concerns in the Fashion Industry: A Twin Transition With the Digital Product Passport. *Business Strategy and the Environment*, 34(7), pp. 9242-9256.

Corallo, Angelo ; Crespino, Anna Maria ; Vecchio, Vito Del ; Lazoi, Mariangela et Marra, Manuela. (2023). Understanding and Defining Dark Data for the Manufacturing Industry. *IEEE Transactions on Engineering Management*, 70(2), pp. 700-712.

Cruz, Estrela Ferreira ; Silva, Pedro ; Serra, Sérgio ; Rodrigues, Rodrigo ; Alves, Marcelo ; Oliveira, João et Rosado Da Cruz, António M. (2025). Machine Learning-Based Data Quality Assessment for the Textile and Clothing Digital Product Passport. *Applied Sciences*, 15(18), pp. 10259.

D'Adamo, Idiano ; Fratocchi, Luciano ; Grosso, Chiara et Tavana, Madjid. (2025). An Integrated Business Strategy for the Twin Transition: Leveraging Digital Product Passports and Circular Economy Models. *Business Strategy and the Environment*, 34(7), pp. 9008-9022.

De Diego, Santiago et Gutiérrez-Aguero, Iván. (2025). Decentralized Digital Product Passport Building Blocks for Enhancing Supply Chain Sovereignty and Circular Economy Practices. *IEEE Access*, 13, pp. 137973-137985.

Do, Namchul. (2017). Integration of design and manufacturing data to support personal manufacturing based on 3D printing services. *The International Journal of Advanced Manufacturing Technology*, 90(9-12), pp. 3761-3773.

Domskienė, Jurgita et Gaidule, Eliza. (2024). An overview of technological challenges in implementing the digital product passport in the textile and clothing industry. *AUTEX Research Journal*, 24(1), pp. 20240002.

Egger, Matthias ; Smith, George Davey et O'Rourke, Keith. (2001). Introduction: rationale, potentials, and promise of systematic reviews. *Systematic reviews in health care: meta-analysis in context*, pp. 1-19.

Fares, Naila ; Islam, Md Shamimul ; Jauhar, Sunil Kumar et Kucukaltan, Berk. (2025). Towards an International Digital Product Passport: The New Paradigm of a Worldwide Circular Economy. *Circular Economy and Sustainability*, 5(6), pp. 5475-5495.

Gieß, Anna et Möller, Frederik. (2025). Exploring the value ecosystem of digital product passports. *Journal of Industrial Ecology*, 29(2), pp. 561-573.

Gonçalves, Anabela ; Leite, Bárbara R. et Silva, Carla. (2025). Sustainability Index in Apparel: A Multicriteria Model Covering Environmental Footprint, Social Impacts, and Durability. *Sustainability*, 17(17), pp. 8004.

Gösken, Christophe. (2025). Fashioning Authenticity: How Digital Product Passports could shape the future of trade mark law. *Journal of Intellectual Property Law and Practice*, 20(2), pp. 105-110.

Hakola, Liisa ; Abedi, Fatemeh ; Nordman, Sirpa ; Smolander, Maria et Paltakari, Jouni. (2025). Smart Tags as Enablers for Digital Product Passports in Circular Electronics Value Chains. *Circular Economy and Sustainability*, 5(3), pp. 2033-2055.

Jeganathan, Kishokanth et Szymkowiak, Andrzej. (2025). Bridging Digital Product Passports and in-store experiences: How augmented reality enhances decision comfort and reuse intentions. *Journal of Retailing and Consumer Services*, 84, pp. 104242.

Jensen, Steffen Foldager ; Kristensen, Jesper Hemdrup ; Adamsen, Sofie ; Christensen, Andreas et Waehrens, Brian Vejrum. (2023). Digital product passports for a circular economy: Data needs for product life cycle decision-making. *Sustainable Production and Consumption*, 37, pp. 242-255.

Jensen, Steffen Foldager ; Kristensen, Jesper Hemdrup ; Christensen, Andreas et Waehrens, Brian Vejrum. (2024). An ecosystem orchestration framework for the design of digital product passports in a circular economy. *Business Strategy and the Environment*, 33(7), pp. 7100-7117.

Kebede, Rahel ; Moscati, Annika ; Tan, He et Johansson, Peter. (2024). A modular ontology modeling approach to developing digital product passports to promote circular economy in the built environment. *Sustainable Production and Consumption*, 48, pp. 248-268.

Kemmner, Felicia-Adina ; Legenvre, Hervé et Hameri, Ari-Pekka. (2025). Sharing data to implement the circular economy: the case of digital product passports. *Industrial Management & Data Systems*, 125(10), pp. 2854-2873.

Kim, Jinyoub ; Moon, Jisang ; Do, Yeji ; Kim, Hayul et Jeong, Jongpil. (2023). Research on a SaaS (Software as a Service)-based Digital Product Passport System Model for the EV Battery Industry Value Chain. *WSEAS TRANSACTIONS ON BUSINESS AND ECONOMICS*, 20, pp. 2466-2475.

King, Melanie R.N. ; Timms, Paul D. et Mountney, Sara. (2023). A proposed universal definition of a Digital Product Passport Ecosystem (DPPE): Worldviews, discrete capabilities, stakeholder requirements and concerns. *Journal of Cleaner Production*, 384, pp. 135538.

Kühn, Maximilian ; Baumann, Michael ; Volz, Friedrich et Stojanovic, Ljiljana. (2025). Digital Product Passport Design Supporting the Circular Economy Based on the Asset Administration Shell. *Sustainability*, 17(3), pp. 969.

Langley, David J. ; Rosca, Eugenia ; Angelopoulos, Marios ; Kamminga, Oscar et Hooijer, Christa. (2023). Orchestrating a smart circular economy: Guiding principles for digital product passports. *Journal of Business Research*, 169, pp. 114259.

Liu, Gang ; Man, Rongjun et Wang, Yanyan. (2021). A Data Management Approach Based on Product Morphology in Product Lifecycle Management. *Processes*, 9(7), pp. 1235.

Mabkhot, Mohammed M. ; Kalawsky, Roy S. et Liaqat, Amer. (2025). Introducing the Manufacturing Digital Passport (MDP): A New Concept for Realising Digital Thread Data Sharing in Aerospace and Complex Manufacturing. *Systems*, 13(8), pp. 700.

Mollahassani, Damun ; Göbel, Jens C. et Juresa, Yannick. (2025). Towards Model-based Definition of Digital Product Passports Supporting Sustainable Smart Product Lifecycles. *Tehnički glasnik*, 19(Si1), pp. 43-48.

Ospital, Pantxika ; Masson, Dimitri ; Beler, Cédrick et Legardeur, Jérémy. (2023). Toward product transparency: communicating traceability information to consumers. *International Journal of Fashion Design, Technology and Education*, 16(2), pp. 186-197.

Panza, Luigi ; Bruno, Giulia et Lombardi, Franco. (2023). Integrating Absolute Sustainability and Social Sustainability in the Digital Product Passport to Promote Industry 5.0. *Sustainability*, 15(16), pp. 12552.

Papile, Flavia et Del Curto, Barbara. (2025). Textile Materials Information for Digital Product Passport Implementation in the Textile and Clothing Ecosystem: A Review on the Role of Raw Fibers in a Substantial Transition. *Sustainability*, 17(19), pp. 8804.

Petrik, Dimitri ; Dzierzawa, Fabian et Warthmann, Kirsten. (2025). A Maturity Model for Digital Product Passports: A Design Science Study. *IEEE Access*, 13, pp. 114575-114594.

Pracucci, Alessandro et Giovanardi, Matteo. (2025). Design of a Sensor-Based Digital Product Passport for Low-Tech Manufacturing: Traceability and Environmental Monitoring in Bio-Block Production. *Sensors*, 25(18), pp. 5653.

Ruismäki, Winnie ; Zu Castell-Rüdenhausen, Malin ; Pohjalainen, Elina et Wahlström, Margareta. (2025). Stakeholder perspectives on digital product passports for construction products. *Cleaner Logistics and Supply Chain*, 17, pp. 100275.

Saenz, José ; Felsch, Torsten ; Walter, Christoph ; König, Tim ; Poenicke, Olaf ; Bayrhammer, Eric ; Vorbröcker, Mathias ; Berndt, Dirk ; Elkmann, Norbert et Arlinghaus, Julia. (2024). Automated disassembly of e-waste—requirements on modeling of processes and product states. *Frontiers in Robotics and AI*, 11, pp. 1303279.

Tranfield, David ; Denyer, David et Smart, Palminder. (2003). Towards a Methodology for Developing Evidence-Informed Management Knowledge by Means of Systematic Review. *British Journal of Management*, 14(3), pp. 207-222.

Vergés, Roger ; Gaspar, Kàtia et Forcada, Núria. (2026). Large-scale analysis of environmental and circularity metrics in construction products through automated data extraction using large language models. *Environmental Impact Assessment Review*, 117, pp. 108243.

Voulgaridis, Konstantinos ; Lagkas, Thomas ; Angelopoulos, Constantinos Marios ; Boulogeorgos, Alexandros-Apostolos A. ; Argyriou, Vasileios et Sarigiannidis, Panagiotis. (2024). Digital product passports as enablers of digital circular economy: a framework based on technological perspective. *Telecommunication Systems*, 85(4), pp. 699-715.

Wan, Paul Keng Fai et Jiang, Shanshan. (2025). Enabling a dynamic information flow in digital product passports during product use phase: A literature review and proposed framework. *Sustainable Production and Consumption*, 54, pp. 362-374.16/j.spc.2025.01.014

Won, Y.; Jang, S.; Kim, H.Y.; Lee, Y.원, 유정 ; 장세윤 ; 김하연 et 이유리. (2025). Digital Product Passports (DPP) in the Textiles, Apparel, and Fashion Industry -Lessons and Recommendations to Realize a Circular Economy-. *Journal of the Korean Society of Clothing and Textiles*, 49(3), pp. 534-550.

Managing Sustainability in a Fashion Supply Chain: The Case of an Italian SME

Sara Ianniello, Livio Cricelli, Serena Strazzullo
University of Naples Federico II, Naples, Italy

Abstract

Sustainability regulations and voluntary standards are increasingly transforming fashion supply chains into information-intensive systems, in which firms are expected to document practices that extend beyond their organisational boundaries. This is particularly relevant in the fashion supply chain, where, although these requirements are usually defined by large focal firms, their implementation often depends on small and medium-sized enterprise (SME) suppliers that collect, interpret, and transmit sustainability information across multiple tiers. However, how SMEs manage sustainability information flows across organisational boundaries remains underexplored. To address this issue, the study draws on knowledge governance literature on supply chains and adopts a single case study based on semi-structured interviews, with an Italian SME operating in a fashion supply chain. Findings show that sustainability information management is mainly driven by customer requests and unfolds as a document-based process of acquisition, internal processing, and dissemination. The SME acts as an intermediary knowledge broker, translating external requirements into operational requests for suppliers, laboratories, and consultants, while sending documentary evidence back to customers. However, the process remains constrained by limited internal verification capabilities and strong dependence on third parties. The analysis identifies a hybrid governance configuration across the fashion supply chain, dominated by contract-based mechanisms, reinforced by market-based pressures, and supported by trust-based coordination, whereas reciprocity-based mechanisms appear marginal. The study contributes to knowledge governance and sustainable supply chain literature by showing how SMEs manage sustainability information under asymmetrical pressure and resource constraints, highlighting both their enabling role and organisational barriers to sustainability implementation. From a managerial perspective, the findings suggest that SMEs should strengthen internal management through simple procedures, clearer allocation of responsibilities, and shared tools for document collection and storage. At the same time, focal firms and policymakers should provide greater support, ensuring that standards, documentary obligations, and reporting tools do not become unsustainable administrative burdens.

Keywords: case study, fashion industry, SME, supply chain, sustainability.

Paper type: Academic Research Paper

1 Introduction

In recent years, sustainability regulations and voluntary standards have progressively expanded firms' responsibilities beyond their own organisational boundaries, requiring them to monitor, document, and report sustainability practices throughout their supply chains (European Commission, 2023). The European Union Corporate Sustainability Reporting Directive (CSRD) reinforces disclosure obligations across the entire value chain, significantly increasing the amount and complexity of sustainability-related information that companies are required to collect, process, and exchange.

At the same time, academic debate has shown a growing relevance of sustainability in global supply chains, encouraging extensive research across sustainability management, supply chain management, and corporate governance. Within this context, the knowledge management literature has conceptualised supply chains as knowledge systems in which competitive advantage depends on resource ownership but also on the ability to create, share, integrate, and apply knowledge across organisational boundaries (Foss & Mahoney, 2010; Zhao et al., 2021).

From this perspective, knowledge governance mechanisms have increasingly gained attention as a perspective for understanding how knowledge processes are structured and coordinated within and across organisations. It refers to the set of formal and informal rules that shape how knowledge is created, shared, accessed, and applied, thus influencing the effectiveness of knowledge-related activities (Foss & Mahoney, 2010; Zhao et al., 2021).

Empirical studies have demonstrated that such mechanisms influence knowledge sharing and integration, ultimately affecting organisational and supply chain performance.

While this distinction provides a useful starting point, it may be insufficient to fully capture the diversity of governance arrangements observed in complex inter-organisational contexts. For this reason, several scholars

propose a more granular classification of knowledge governance mechanisms into four main types of mechanisms (Zhao et al., 2021): trust, market, reciprocity, and contract-based.

However, this stream of research has largely focused on focal firms' innovation outcomes, with limited attention to sustainability enforcement along supply chain.

Only recently have scholars begun to acknowledge that implementing sustainability is inherently knowledge-intensive, requiring continuous interpretation, translation, and validation of heterogeneous information among multiple actors (Miandar et al., 2024). Gong et al. (2018) show that sustainability in multi-tier supply chains is learned through processes of resource orchestration involving focal firms and supply chain partners, and that it requires deliberate efforts to mobilise resources, coordinate actors, and support the acquisition and internalisation of sustainability. Similarly, Gualandris et al. (2015) suggest that sustainability in supply chains is not simply a matter of compliance, but also a matter of organising information in processes embedded in everyday organisational practices.

This issue is particularly significant in fashion supply chains, which are typically organised as multi-tier and geographically dispersed production networks. In the European fashion supply chain, apparel production is frequently supported by networks of specialised small and medium-sized enterprises (SMEs) involved in sourcing, processing, subcontracting, and manufacturing activities (European Commission, 2023). Specifically, the European fashion value chains rely on a dense base of Italian suppliers, while the broader textiles ecosystem is overwhelmingly composed of SMEs. Within this context, focal firms coordinate design, branding, and market access, while a wide range of manufacturing and sourcing activities are carried out by legally independent and often small actors (Arimany Serrat et al., 2025; Singhal, 2023). Thus, sustainability requirements are often operationalised by SMEs that must transmit them across the supply chain. Existing studies have shown that first-tier suppliers play an important coordinative and interpretive role in cascading sustainability demands along multi-tier supply chains (Jia et al., 2021); yet little is known about how SMEs manage sustainability and which governance mechanisms emerge with the other actors involved in processing it. From a theoretical perspective, this raises important questions for the knowledge governance literature. Prior studies on knowledge governance mechanisms have mainly examined how such mechanisms support knowledge exchange, integration and supply chain performance (Zhao et al., 2021). Less attention has been paid to how these mechanisms operate when SMEs are required to manage sustainability-related information under asymmetric customer pressure, limited internal resources and dependence on external actors (Abdul Basit et al., 2024; Nath et al., 2024). As a result, knowledge governance mechanisms may operate differently. Accordingly, this study contributes to the knowledge governance literature and addresses the following research questions:

- ***RQ1****: How does an SME operating in a European fashion supply chain manage the sustainability information?*
- ***RQ2****: How do knowledge governance mechanisms along the European fashion supply chain shape sustainability information management within an SME?*

To answer these research questions, the authors adopt an exploratory single case study approach based on semi-structured interviews, designed for an Italian SME operating in a European fashion supply chain. The aim is to understand how the SME manages sustainability information in response to external demands, how different governance mechanisms shape the process across the supply chain actors, and what drivers and barriers emerge in this process.

2 Theoretical Perspective

2.1 Knowledge governance mechanisms

Durst & Edvardsson (2012) define knowledge management as the set of processes and structures through which firms support knowledge-related activities such as transfer, storage, and creation. Within this broad stream, a more specific line of investigation has developed around knowledge governance. This perspective examines the mechanisms through which such processes are coordinated and regulated. As Zhao et al. (2021) explain, a common distinction in the literature differentiates between formal and informal knowledge governance mechanisms. Formal mechanisms rely on codified structures and managerial tools, such as performance evaluation systems, incentives, training programs, and reward schemes, which are designed to motivate individuals to engage in knowledge creation, sharing, and transfer activities. In contrast, informal mechanisms are rooted in social and relational dimensions, including trust, shared norms, and teamwork, which facilitate knowledge exchange through interpersonal relationships and mutual understanding.

Compared to the broader distinction between formal and informal mechanisms, there is another approach that divides mechanisms into four categories (Zhao et al., 2021):

1. Trust-based mechanisms emphasise the role of mutual confidence in enabling knowledge exchange. Trust reduces perceived risks associated with opportunistic behaviour and facilitates the transfer of tacit knowledge, which is often difficult to codify. Prior research highlights several factors that contribute to the development of trust, including prior relationships, third-party referrals, open communication, and geographical proximity.
2. Market-based mechanisms, on the other hand, rely on price signals and economic incentives to regulate knowledge exchange. In this case, knowledge is treated as a tradable resource, and its transfer occurs through market transactions based on supply and demand conditions.
3. Reciprocity-based mechanisms are grounded in mutual exchange relationships, where actors share knowledge with the expectation of future returns. Drawing on the notion of reciprocity as "giving and receiving according to need", this mechanism fosters stable and long-term relationships, which in turn facilitate the sharing of valuable knowledge and resources. Such relationships are particularly important in networked environments, where sustained interactions enhance collaboration and knowledge diffusion.
4. Contract-based mechanisms rely on formal agreements that define roles, responsibilities, and obligations among partners. Contracts function both as coordination tools and as safeguards against opportunistic behaviour and uncertainty. In addition, they allow adaptation to changing conditions and can support learning processes by structuring interactions and clarifying expectations.

This approach enables the identification of specific governance practices and their role in facilitating interactions among actors. For these reasons, it has been widely used to explain how knowledge exchange is regulated in inter-organisational contexts, particularly in supply chains characterised by distributed knowledge, heterogeneous actors, and coordination challenges (Foss & Mahoney, 2010). This is particularly relevant in the fashion supply chain, when lots of SMEs are required to process a wide range of information, with limited internal resources and high dependence on external actors (Sartori et al., 2020). Compared to large firms, SMEs typically rely on less formalised approaches, as financial, human, and technological constraints limit the adoption of structured systems and dedicated functions (Abdul Basit et al., 2024; Nath et al., 2024; Sartori et al., 2020). Related studies on fashion supply chains show that sustainability implementation depends on the ability to coordinate multiple actors across tiers. Karaosman et al. (2020) demonstrate that this coordination is built through a combination of formal and relational practices, including supplier engagement, long-term trading relationships, and the mobilisation of technical and relational resources. However, they also show that these practices do not automatically ensure full transparency, since end-to-end visibility remains difficult to achieve in fragmented fashion supply chains. Some scholars extend this view: sustainability-oriented innovation emerges through the ability of a focal SME to monitor its external environment, involve relevant partners, combine internal and external resources, and activate capabilities across the network (Ermini et al., 2024).

2.2 Conceptual model

In the Italian context, the fashion industry is characterised by a dense network of SMEs that operate within European supply chains. Within these networks, SMEs are increasingly required to align with the sustainability standards, reporting practices, and information requirements imposed by focal firms. This growing external pressure suggests that, even in traditionally informal organisational settings, SMEs may progressively adopt more structured and formalised governance mechanisms within the supply chain to manage sustainability information.

As a result, contract- and market-based mechanisms, which have generally been less visible, may become more relevant when considering the entire chain currently.

Building on this perspective, the present study adopts a knowledge governance approach as the main theoretical lens. While Zhao et al. (2021) examine knowledge governance mechanisms mainly in relation to supply chain performance, this study uses the same categorisation to investigate how sustainability information is managed across supply chain actors. In line with Gualandris et al. (2015), sustainability information management is conceptualised here as a process characterised by three sequential phases: acquisition, internal processing, and dissemination across the supply chain. From this perspective, the study investigates how different governance mechanisms between the SME and the other actors influence how an SME acquires, interprets, and shares sustainability-related information. Figure 1 presents the conceptual framework guiding the analysis.

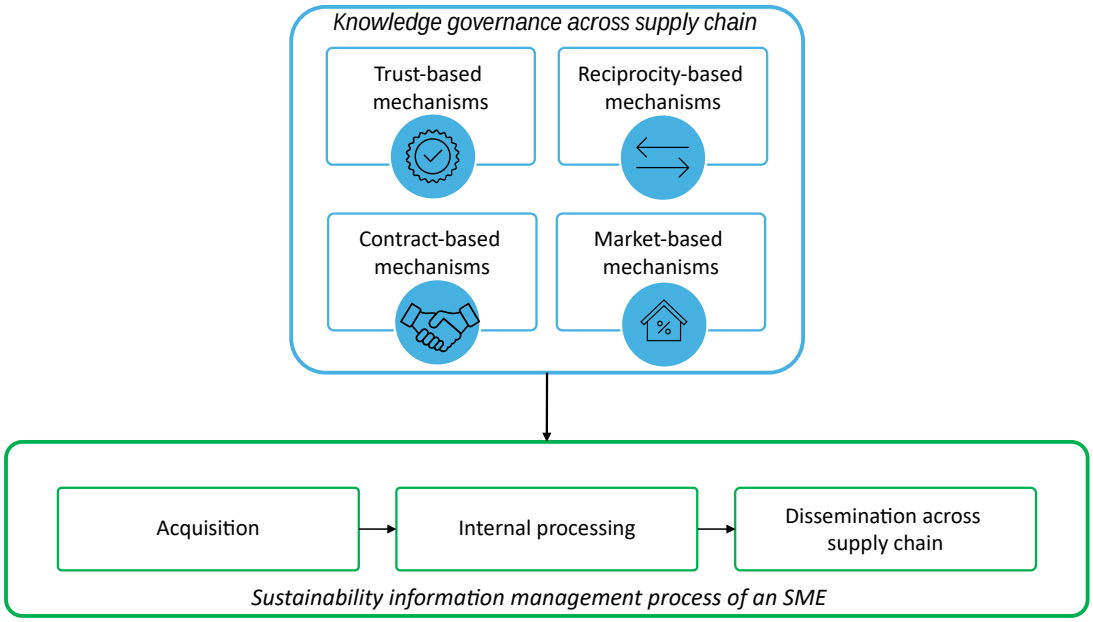


Figure 1: Conceptual framework adapted from Zhao et al. (2021).

3 Method

This study adopts an exploratory single case study, based on semi-structured interviews, widely recognised as an effective method for exploring knowledge processes and governance mechanisms. A single case study is particularly appropriate when the research aims to answer "how" and "why" questions and to explore contemporary phenomena within their natural context. Moreover, it allows for an in-depth investigation of organisational processes that are difficult to observe through large-scale quantitative methods.

Case selection

The study focuses on a single Italian SME operating within a European fashion supply chain. The firm was selected through purposeful sampling because it represents an information-intensive context. Specifically, the company coordinates relationships between international fashion brands and upstream suppliers in the manufacture of men's trousers. This positioning makes it a particularly suitable case for investigating how the sustainability information process is managed along the supply chain.

Data collection

The interviews involved 5 organisational actors whose expertise areas are detailed in the following Table 1:

Table 1: Characteristics of the interviewed experts.

Expert	Organizational role	Main expertise
1	Production	Materials, fabrics, product tests, supplier interaction, laboratory testing
2	Administration	Administrative documentation, contracts, supplier regularity, formal declarations
3	Ownership and commercial management	Client relationships, market pressures, supplier coordination, responsibility allocation
4	Commercial and customer interface	Client requests, traceability, certifications, supplier documentation, information flows
5	Commercial and operational support	Customer requests, supplier communication, product information

The interview protocol was designed to explore all the dimensions identified in the conceptual model:

1. how sustainability information is collected and received from external stakeholders;
2. how such information is internally processed, interpreted, and validated;
3. how it is shared and disseminated across supply chain actors;
4. what governance mechanisms are implemented along the supply chain;
5. relations between the process and the governance mechanisms adopted, complemented by a matrix;

3.1 Data analysis

First, the empirical material was analysed using a thematic analysis approach. Following Saldaña (2013), interview data were analysed through a two-step coding process. First, relevant extracts from the transcripts were condensed into brief descriptive labels, generating initial nodes and preliminary sub-themes. These were then reviewed and grouped according to conceptual similarity. The analysis followed three coding steps: open coding, to inductively identify categories emerging from participants' accounts; axial coding, to explore relationships among codes; and selective coding, to refine higher-order themes capturing the core patterns of the data. An example of the coding process is provided in Table 2.

Table 2: Example of the coding process.

Expert	Interviews sentences	Open code	Axial code	Selective code
1	*"Clients request certifications, chemical tests, fabric conformity documents and laboratory analyses, such as REACH, OEKO-TEX and product-specific tests"*	Documentary and technical sustainability requirements	External pressure and documentary formalization	Sustainability information management as externally driven compliance work

The second stage of the data analysis involved the expert-based construction of an impact matrix aimed at assessing the influence of each knowledge governance mechanism on the different phases of the sustainability information management process. Experts were asked to assign a score from 1 to 3, where 1 indicated a weak impact, 2 a moderate impact, and 3 a strong impact. A weak impact was assigned when the mechanism did not influence how sustainability information was managed. A moderate impact indicated that the mechanism was present only marginally or occasionally. A strong impact was assigned when the mechanism made a relevant contribution to the process.

4 Results

4.1 Thematic analysis

The coding process identified four main themes that explain how sustainability information management unfolds across internal activities and supply chain relationships. These themes are discussed below.

4.1.1 Sustainability information management process

Regarding the first research question, the empirical evidence shows that sustainability information management, rather than emerging as a fully internalised managerial process, takes the form of a reactive and document-based activity through which the SME receives, interprets, collects and transmits sustainability requests of downstream actors.

In particular, the company is positioned between downstream customers and upstream suppliers, laboratories, consultants and production partners. Its role consists of receiving customer requirements, translating them into operational and documentary requests, mobilising the relevant actors and transmitting the collected evidence back to customers. Although the process is analytically divided into acquisition, internal processing and dissemination, the empirical evidence shows that these phases are recursive rather than strictly sequential. Specifically, requirements are received from clients, usually through emails, order-related documents, specifications, questionnaires, or contractual documents. These requirements are not limited to environmental issues, nor do they only concern product certification. They include product-related tests, chemical compliance, material certifications, traceability information, supplier declarations, social and ethical commitments, administrative documents, contractual obligations and, in some cases, customer-specific insurance requirements. The type of information requested varies depending on the client, the product, and the extent to which the SME directly manages the supply of fabrics and accessories. When customers indicate already approved suppliers, the SME assumes that such suppliers are likely to possess the required certifications. The requirements identified through interviews have been grouped in Table 3:

Table 3: Main sustainability requirements and information that SME manages.

Category	SME sustainability requirements
Product safety, chemical and environmental compliance	REACH compliance for fabrics
	Oeko-Tex certification for fabrics/materials
	Tests on fabrics, accessories, dyes, prints or treatments
	PFAS tests
	Evidence of absence or limited presence of harmful substances
	Laboratory test reports issued by certified or recognized bodies
Durability and product-performance requirements	Cross-staining test
	Colour fastness test
	Light fastness test
	Anti-pilling test
	Antistatic treatment for synthetic fabrics
Materials and product origin	Information on raw material origin
	Information on materials used in garments
	Certificates of origin
	Information on upstream suppliers, sub-suppliers or raw material producers
	Traceability information on the production chain
Social and ethical requirements	Ethical commitments regarding outsourced production activities
	Commitments concerning workers' treatment and labour-related standards
	Evidence of compliance with customer-defined social standards
	Codes of conduct requested by customers
Supplier responsibility and documentary assurance	Updated supplier sustainability certifications
	Signed declarations by suppliers assuming responsibility for the truthfulness of sustainability-related information
	Declarations of conformity concerning materials, products or production standards
Audit and inspection requirements	Customer audits on supply-chain conduct or sustainability-related requirements
	Physical inspections at suppliers, laundries, "façoniste", or foreign production sites
	Evidence of production processes when required by customers for sustainability or traceability purposes

Following, the SME collects the required documents from suppliers, external laboratories, consultants, accountants or other supply chain actors. Importantly, the empirical evidence does not support the existence of a systematic internal verification phase. The SME does not usually have internal tools or dedicated competencies to technically verify the sustainability information received. Instead, the firm relies on external sources of assurance. Finally, the information is transmitted back to the client, generally through email, PEC or standard documentation.

4.1.2 Sustainability information management as externally driven compliance work

The second theme shows that sustainability does not emerge primarily as a personal strategy, but as a set of documentary and technical obligations imposed by downstream actors. Interviewees report that customers increasingly request evidence of compliance. The owner confirms that medium-high brands request several requirements, creating internal organisational work to respond to these demands. As a result, the SME is indirectly exposed to sustainability regulation through the imposition by focal firms, while remaining only marginally involved in the definition of the standards it is expected to implement. This creates a cascading accountability structure, in which the focal firm becomes responsible not only for its own practices, but also for demonstrating the compliance of suppliers, subcontractors and material providers. Such a dynamic is particularly challenging for a small firm, as it must respond to complex and evolving requirements without necessarily having the internal resources, expertise or bargaining power of larger organisations.

4.1.3 Hybrid governance: formal requirements managed through informal coordination

The interviews point to a hybrid form of governance in the management of sustainability information. Formal and market-based mechanisms have become increasingly relevant, as they exert pressure on the SME to comply with external expectations to maintain commercial relationships and access to important clients. However, formal mechanisms do not fully replace relational forms of governance. Trust-based relationships with long-standing clients and suppliers remain essential for facilitating communication, accelerating information

exchange and reducing uncertainty in situations where documentation is incomplete or difficult to obtain. This is especially problematic when information concerns upstream actors, indirect suppliers or the origin of materials, where transparency is more difficult to achieve. Moreover, since the SME lacks dedicated internal tools and technical competencies to produce and then verify this information directly, the internal processing phase remains dependent on third-party declarations and documentation.

Therefore, the sustainability information management process relies on the coexistence of contractual obligations, market pressures and relational trust. This hybridity reflects the practical need to combine formal compliance with informal coordination to manage complexity across the fashion supply chain.

4.1.4 Learning-by-doing under resource and capability constraints

The last theme captures the learning process through which the SME gradually develops practical knowledge about sustainability information management. Interviewees explain that certifications such as OEKO-TEX and REACH were initially difficult to understand; over time, however, the company learned which actors to contact, how to request documents and how to obtain information more quickly. This learning remains fragile because it depends on accumulated experience, external consultants, suppliers and a limited number of internal actors. The owner emphasises that a small firm would need legal, accounting and technical expertise to manage all customer requirements precisely, but such competences are mostly external and limited by time and cost. Overall, low competences, limited resources and lack of dedicated personnel remain key barriers.

4.2 Relational model

The thematic analysis was complemented through an expert-based matrix in which each interviewee assessed the perceived impact of four knowledge governance mechanisms, adopted along the supply chain, on the three phases of sustainability information management. The results show that contract-based mechanisms have the strongest overall impact, followed by market-based and trust-based mechanisms, while reciprocity-based mechanisms remain weak and marginal, consistent with the qualitative evidence. Average scores across experts were then synthesised into the relational model presented in Figure 2:

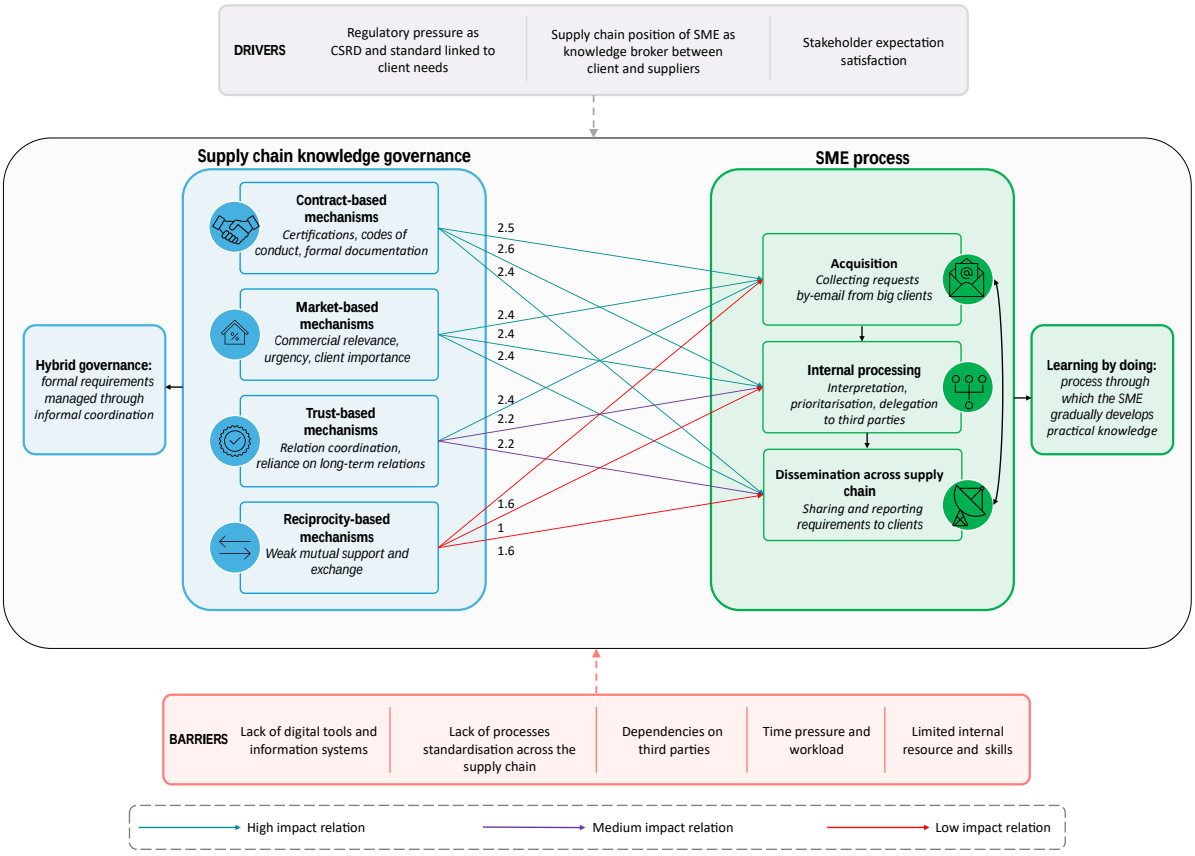


Figure 2: Relational model.

The model suggests that the sustainability information process is primarily activated by external drivers, including regulatory pressure, client requirements, and stakeholder expectations, consistent with Gualandris et al. (2015). The bidirectional arrow between acquisition and dissemination indicates that sustainability information management is not a linear process. According to the first theme, information circulates recursively from clients to the SME, from the SME to third parties, and back again to the SME and clients.

The central pattern emerging from the model is the hybrid governance configuration. Formal and informal mechanisms coexist, but they do not contribute equally to the process. Contract-based mechanisms structure what information must be collected, how it is validated, and how it is transmitted, through certifications, codes of conduct, test reports, supplier declarations, and client documentation. Market-based mechanisms further reinforce this formal orientation, as client importance, commercial urgency, and order relevance influence the prioritisation of sustainability requests.

Trust-based mechanisms play a supporting role: long-term relationships with clients and informal communication with suppliers help the SME obtain missing information, clarify ambiguous requests, and manage clients, but they do not replace formal documentation. Reciprocity-based mechanisms, instead, remain weak, suggesting that the process is more of a cascade of requests than a collaborative construction of knowledge.

Compared with studies that emphasise the informal nature of knowledge management in the fashion supply chain, informal practices are present (Ermini et al., 2024; Karaosman et al., 2020), but they are mainly activated to make formal requirements workable due to limited digital tools, incomplete procedures, uncertain supplier responses, time pressure, and resource constraints. This pattern is consistent with Ermini et al. (2024), who underline the importance of informal coordination mechanisms in sustainability implementation, but our findings add that these practices emerge as a pragmatic response to the incompleteness of formal systems. Over time, these constraints lead the SME to develop practical capabilities through learning by doing. In this sense, the role of knowledge governance differs from the innovation perspective often adopted in prior studies, such as Zhao et al. (2021): here, governance mechanisms aim to enable operational coordination, information reliability, and the circulation of sustainability knowledge across interdependent actors in the supply chain.

5 Conclusions

This study investigated how sustainability information is acquired, processed and disseminated within an Italian SME operating in a European fashion supply chain, and how knowledge governance mechanisms along the supply chain shape this process. Regarding the first research question, the findings show that sustainability information management is mainly activated by downstream customer requirements and unfolds as a document-based and recursive process. As a result, the SME performs an intermediary knowledge-brokering role, which interprets and coordinates heterogeneous sustainability requirements across multiple actors.

The research contributes to the knowledge governance literature by showing that, although sustainability information management within the SME remains partly informal and experience-based, its inter-organisational configuration is strongly shaped by contract-based and market-based mechanisms imposed along the supply chain. Thus, for the second research question, trust-based mechanisms do not replace formal requirements, but support their practical management when documentation is incomplete, ambiguous or difficult to obtain. Contract-based mechanisms define the documentary obligations, standards and responsibilities attached to sustainability requirements, while market-based mechanisms reinforce compliance through customer pressure, commercial urgency and the need to maintain access to valuable business relationships. By contrast, reciprocity-based mechanisms appear relatively weak, suggesting that sustainability information management is more of a cascading process of requirements, responsibilities and documentary evidence. These findings also extend the sustainable supply chain literature by highlighting the hidden organisational work performed by an SME in the implementation of sustainability standards. Therefore, the case shows that SME acts as critical but vulnerable nodes in sustainability governance: it absorbs administrative burdens, uncertainty and responsibility without always having adequate resources, expertise or digital tools.

From a managerial perspective, the findings suggest that the Italian SME should strengthen the internal organisation of sustainability information through simple but systematic practices, such as shared document repositories, standardised checklists, clearer responsibility allocation, and basic procedures for collecting, storing and verifying sustainability-related documents. These tools would not eliminate the need for external certifications or supplier declarations, but they could reduce dependence on individual experience and improve the reliability and traceability of information flows. At the same time, focal firms should recognise that sustainability requirements generate significant organisational work for SME suppliers. Similarly, policymakers

should consider the resource constraints of SMEs when designing reporting obligations and support mechanisms.

Finally, the study is subject to some limitations. First, it is based on a single case study and therefore does not aim at statistical generalisation. Second, the analysis focuses on the perspective of the SME, while future research could include clients, suppliers, laboratories and consultants to reconstruct the full multi-tier sustainability information flow. Third, the expert-based matrix captures perceived impacts of governance mechanisms and could be further developed or validated through comparative research. Moreover, future studies could examine how digital tools, shared platforms and standardised processes moderate the relationship between knowledge governance mechanisms and sustainability information management.

6 Acknowledgements

The authors acknowledge the valuable participation and support of the Italian SME in this research.

7 Ethics declaration

No ethical clearance was required for this study.

8 AI declaration

The authors declare the use of ChatGPT to support language refinement.

References

Abdul Basit, S., Gharleghi, B., Batool, K., Hassan, S. S., Jahanshahi, A. A., & Kliem, M. E. (2024). Review of enablers and barriers of sustainable business practices in SMEs. *Journal of Economy and Technology*, *2*, 79–94. https://doi.org/10.1016/j.ject.2024.03.005

Arimany Serrat, N., Arribas-Ibar, M., & Erdoğan, G. (2025). Fast Fashion Sector: Business Models, Supply Chains, and European Sustainability Standards. *Systems*, *13*(6). https://doi.org/10.3390/systems13060405

Durst, S., & Edvardsson, I. R. (2012). Knowledge management in SMEs: A literature review. In *Journal of Knowledge Management* (Vol. 16, Number 6, pp. 879–903). https://doi.org/10.1108/13673271211276173

Ermini, C., Visintin, F., & Boffelli, A. (2024). Understanding supply chain orchestration mechanisms to achieve sustainability-oriented innovation in the textile and fashion industry. *Sustainable Production and Consumption*, *49*, 415–430. https://doi.org/10.1016/j.spc.2024.07.008

European Commission. (2023). *Corporate Sustainability Reporting*.

Foss, N. J., & Mahoney, J. T. (2010). Exploring knowledge governance. In *Organization Science*, *Strategic Entrepreneurship Journal*, *Strategic Management Journal and Strategic* (Vol. 2, Number 3).

Gong, Y., Jia, F., Brown, S., & Koh, L. (2018). Supply chain learning of sustainability in multi-tier supply chains: A resource orchestration perspective. *International Journal of Operations and Production Management*, *38*(4), 1061–1090. https://doi.org/10.1108/IJOPM-05-2017-0306

Gualandris, J., Klassen, R. D., Vachon, S., & Kalchschmidt, M. (2015). Sustainable evaluation and verification in supply chains: Aligning and leveraging accountability to stakeholders. *Journal of Operations Management*, *38*, 1–13. https://doi.org/10.1016/j.jom.2015.06.002

Jia, M., Stevenson, M., & Hendry, L. C. (2021). The boundary-spanning role of first-tier suppliers in sustainability-oriented supplier development initiatives. *International Journal of Operations and Production Management*, *41*(11), 1633–1659. https://doi.org/10.1108/IJOPM-12-2020-0856

Karaosman, H., Perry, P., Brun, A., & Morales-Alonso, G. (2020). Behind the runway: Extending sustainability in luxury fashion supply chains. *Journal of Business Research*, *117*, 652–663. https://doi.org/10.1016/j.jbusres.2018.09.017

Miandar, T., E. Johnsen, T., & Caniato, F. (2024). The role of purchasing and supply management in diffusing sustainability in supply networks: A systematic literature review. *Business Ethics, the Environment and Responsibility*, *33*(3), 505–522. https://doi.org/10.1111/beer.12622

Nath, S. D., Eweje, G., & Barua, S. (2024). Drivers and barriers for implementing social sustainability in supply chains: a qualitative investigation of a developing country's multi-tier suppliers. *International Journal of Logistics Management*, *35*(4), 1332–1367. https://doi.org/10.1108/IJLM-03-2022-0121

Saldaña, J. (2013). *The Coding Manual for Qualitative Researchers*. www.sagepublications.com

Sartori, J. T., Frederico, G. F., do Carmo Duarte Freitas, M., & Mendes Júnior, R. (2020). Specificities of SMEs relevant to knowledge management: a systematic literature review "Specificities of SMEs relevant to knowledge management: a systematic literature review." In *Int. J. Business Excellence* (Vol. 22, Number 1).

Singhal, N. (2023). Stakeholders sustainable development goals (SDGs) prioritization. *Business Strategy and Development*, *6*(4), 986–990. https://doi.org/10.1002/bsd2.292

Zhao, G., Hormazabal, J. H., Elgueta, S., Manzur, J. P., Liu, S., Chen, H., Lopez, C., Kasturiratne, D., & Chen, X. (2021). The impact of knowledge governance mechanisms on supply chain performance: empirical evidence from the agri-food industry. *Production Planning and Control*, *32*(15), 1313–1336. https://doi.org/10.1080/09537287.2020.1809023

Competence Model for Collaborative Workplaces in Industry 5.0: Initial Insight

Tadej Crnjac[1], Robert Ojsteršek1, Marta Aranyossy[2], Zlatko Nedelko[1,2]
[1] University of Maribor, Maribor, Slovenia
[2] Corvinus University of Budapest, Budapest, Hungary

Abstract

Industry 5.0 emphasizes the synergy between humans and technology, where employees' competencies are critical for the effective implementation of collaborative workplaces. In this context, collaborative robots play a key role in modernizing organizational processes by increasing flexibility, usability, and reprogrammability (Segura et al., 2021; Ojsteršek et al., 2024). Research on human–robot collaboration (HRC), grounded in task analysis and ISO 10218-2 safety guidelines (International Organization for Standardization, 2011), highlights the need to align task requirements with human capabilities and competency profiles. Effective HRC requires not only technical knowledge but also cognitive and psychosocial competencies (Wilson & Salas, 2010; Villani et al., 2018; Salas and Rosen, 2020). Existing models typically distinguish four competency clusters: technical, safety, cognitive, and psychosocial. However, none provides a fully comprehensive framework for collaborative robotics (Werning et al., 2021; Simone et al., 2022; Wang et al., 2023; Potocan et al., 2024). This study addresses the gap by proposing a methodology for predicting competency weights, enabling structured learning approaches and supporting future empirical validation. Safety competencies are treated as embedded within technical, cognitive, and psychosocial dimensions, reflecting their integrative role in ISO-based HRC environments. The article develops an enhanced competency model based on task and job analysis aligned with ISO standards and applies the Analytic Hierarchy Process (AHP) to determine competency weights according to their impact on safety, efficiency, and work quality. The model integrates technical, cognitive, and psychosocial competencies, as well as ethical aspects, and emphasizes continuous professional development. To identify competencies, we combined a literature review with the Delphi method, collecting expert opinions through iterative rounds. This resulted in a comprehensive set of existing and additional competencies enabling safe and adaptive performance in collaborative environments. Employees were categorized into three groups: experts (Group A), students (Group B), and laypersons (Group C), reflecting different levels of prior knowledge and experience. Using ISO-based analysis, base competency weights were derived as follows: technical (0.30), cognitive (0.40), and psychosocial (0.30) (Romero et al., 2016; 2022). The model further combines AHP with TOPSIS, a robust multi-criteria decision-making approach (Opricovic & Tzeng, 2004). The main contribution lies in the methodological and normative development of a competency model for Industry 5.0 collaborative environments. The proposed framework establishes a foundation for empirical validation and practical application in personnel selection, training design, and task adaptation across different user groups.

Keywords: Competency Model, Industry 5.0, Human-Robot Collaboration, Human-Centered Manufacturing.

1 Introduction

The implementation of Industry 5.0 represents a major technological shift that transforms business environments and managerial paradigms (Richa Saxena & Arya, 2023). Within this context, collaborative robots play a central role by enhancing flexibility, usability, and reprogrammability in production systems (Segura et al., 2021; Ojsteršek et al., 2024). Their integration requires technical, cognitive, and psychosocial competencies to ensure safe and effective human–robot collaboration (Villani et al., 2018). However, literature highlights significant gaps in existing models. Several studies (Simone et al., 2022; Werning et al., 2021; Wang et al., 2023) emphasize insufficient attention to cognitive and psychosocial competencies, safety, ethical aspects, and continuous learning, as current approaches remain overly focused on technical skills (Potocan et al., 2024). This paper addresses this gap by proposing a methodology for predicting competency weights in collaborative workplaces, enabling structured learning design and supporting future empirical validation. The study develops an enhanced competency model that combines ISO-based task and job analysis with the Analytic Hierarchy Process (AHP) to determine the relative importance of technical, cognitive, and psychosocial competencies, including safety and ethical aspects, while emphasizing continuous professional development.

2 Existing and Additional Competencies

Through an in-depth analysis of competencies and requirements for the successful performance of defined tasks, while considering the work environment and safety standards in accordance with ISO 10218-2 guidelines (International Organization for Standardization, 2011), many publications on competency categorization in smart manufacturing (Frey & Osborne, 2017; Bughin et al., 2018) follow the study of Wilson and Salas (2010).

The authors identified four main competency clusters forming the basis of competency models in smart manufacturing: technical, safety, cognitive, and psychosocial competencies. The following chapters present the methodology for predicting the weights of technical, cognitive, and psychosocial competencies using the ISO analytical model, the AHP method, and the TOPSIS method (Technique for Order Preference by Similarity to Ideal Solution). The methodological framework begins with a theoretical review of competencies and the development of a comprehensive competency set. Competencies are then assigned to groups A, B, and C according to predefined criteria and importance. In the next phase, the AHP method is applied to determine competency weights through pairwise comparison, while the TOPSIS method is used to identify the optimal alternative based on the relative closeness index. Finally, the framework predicts the most suitable group for working with a collaborative robot and validates the model through the Delphi method and expert interviews.

2.1 Existing competencies

Existing competencies include the knowledge, skills, and abilities acquired through education and experience that enable individuals to manage complex systems and ensure safe and productive environments. Technical competencies remain the foundation of smart manufacturing and include knowledge of mechatronics, data analysis for production optimization and fault detection (Bahrin et al., 2016), and understanding of safety standards such as ISO 10218-1 (International Organization for Standardization, 2011). Alongside technical skills, cognitive competencies such as problem-solving, critical thinking, and adaptability are essential due to rapid technological change (Frank et al., 2019), while psychosocial competencies, including teamwork, communication, and stress resilience, are increasingly important in human-centered production environments (Spöttl & Windelband, 2021). Their integration, combined with continuous learning, is crucial for successful adaptation to digital transformation (Liu et al., 2024). Furthermore, traditional models often lack mechanisms for continuous competency development in rapidly evolving production systems (Liu et al., 2022).

2.2 Additional competencies

For the successful integration of collaborative robots into the production process, individuals must develop additional competencies to perform effectively in collaborative environments (Simone et al., 2022). In defining the key competencies of operators working with collaborative robots, this paper assumes that workplaces already include clearly defined tasks and focuses on the additional competencies required for safe and efficient human–robot interaction. Particular attention is given to ISO 10218-1, ISO 10218-2, and ISO/TS 15066 (International Organization for Standardization, 2016), which represent the main international standards in the field of collaborative robotics (Matheson et al., 2019). ISO 10218-1 defines safety requirements for the design and manufacture of industrial robots, including protective functions such as speed limitations, power and force monitoring, and emergency stop systems. ISO 10218-2 complements these requirements by focusing on the safe integration of robots into the work environment through protective barriers, safety sensors, and risk-reduction procedures. ISO/TS 15066 specifically addresses collaborative robots and safety in shared workspaces by defining threshold values for safe human–robot interaction and providing guidelines for risk assessment and permissible loads on different parts of the human body.

2.3 Assessment of additional competencies

To verify the validity of the proposed and additional competencies, we applied the Delphi method, systematically collecting expert opinions through multiple video-based rounds. This structured approach, based on iterative, anonymous surveys with feedback, is widely used to align expert views (Skulmoski et al., 2007) and is suitable for developing competency models in advanced technological contexts. The process included three rounds. In the first, experts were interviewed using a semi-structured questionnaire to evaluate technical, cognitive, and psychosocial competencies. Qualitative content analysis with thematic coding showed strong consensus on technical competencies (e.g., HMI, safety protocols) and cognitive competencies (e.g., problem-solving, spatial orientation), while psychosocial competencies (e.g., teamwork, stress resilience) were also identified as essential. Based on these findings, the second round introduced structured Likert-scale questions to assess competency relevance, identify gaps, and highlight challenges in human–robot collaboration, contributing to a more comprehensive validation of the model. In the third round, experts received group feedback and could revise their assessments, increasing consensus—especially for competencies with higher variability. Strong agreement was reached for most technical and cognitive competencies, while psychosocial competencies remained more diverse. The outcome was a validated list of key additional competencies, incorporated into the further development of the comprehensive competency model.

3 Comprehensive Competency Model

The process of integrating existing smart manufacturing competencies with additional competencies related to collaborative robots requires a systematic approach. This includes the analysis of production processes, assessment of job roles, consultation with experts, identification of additional competencies, and the integration of comprehensive competencies into a unified competency framework. The first step in this process is production analysis, where key processes in industrial environments—already using or planning to implement collaborative robots—are examined. This phase involves reviewing existing research, industry standards, and competency frameworks defined by ISO standards. It is followed by an analysis of job roles, technological requirements, and strategic objectives of the companies, focusing on how the introduction of collaborative robots affects task requirements. This phase includes a systematic evaluation of changes in technical knowledge, cognitive abilities, and psychosocial skills required by employees to work effectively in a collaborative robot environment (Dabić et al., 2021).

3.1 Formulation of comprehensive competencies

Existing and additional competencies together form a comprehensive framework enabling effective, safe, and adaptive work in collaborative environments. This study examines how technical, cognitive, and psychosocial competencies influence productivity in human–robot collaboration and their role in selecting suitable individuals. These dimensions must therefore be integrated into a unified competency profile to ensure successful performance (Welfare et al., 2019; Roychoudhury, 2023). Participants were divided into three groups: experts (A), students (B), and laypersons (C), reflecting different levels of prior knowledge and experience. Technical competencies were assessed differently across groups, while cognitive and psychosocial competencies were evaluated uniformly. Technical competencies represent the foundation for working with collaborative robots and include knowledge of programming, artificial intelligence, sensors, and automation (Huang & Wang, 2023). Experts require advanced understanding of robot operations (Siciliano & Khatib, 2016; Villani et al., 2018), whereas students possess mainly theoretical and basic practical knowledge (Smith & Peterson, 2019). Laypersons, in contrast, require only minimal technical skills. Cognitive competencies support problem-solving, decision-making, and adaptation in dynamic environments (Frank et al., 2019), while psychosocial competencies such as teamwork, communication, and stress management directly influence performance and safety in human-centered production systems (Le Blanc et al., 2020; Parker & Griffin, 2022).

3.2 Validation and generalization of the comprehensive competency model

The comprehensive competency model, integrating technical, cognitive, and psychosocial competencies, is designed to ensure effective, safe, and adaptable work in collaborative environments. Its scientific credibility was assessed through expert validation involving specialists in industrial robotics, ergonomics, work psychology, and human resource development. Following Thürer et al. (2022), the evaluation focused on the adequacy of competencies, clarity of definitions, and relevance for real industrial environments. Although primarily developed for collaborative robots, the model can be extended to other smart manufacturing systems. In Industry 5.0, employees increasingly work with technologies such as smart exoskeletons, digital twins, flexible manufacturing systems, AR/VR, and autonomous mobile robots (AMRs), all requiring varying combinations of technical, cognitive, and psychosocial competencies. The model's generalization is based on key factors, including the degree of human–machine interaction, collaboration requirements, technological complexity, and psychological resilience. Table 1 presents a comparative assessment of competency requirements across selected technologies.

Table 1: Comparative applicability of the competency model across smart technologies

Technology	Technical Competencies	Cognitive Competencies	Psychosocial Competencies
Collaborative robots	5	5	5
Smart exoskeletons	3	4	5
Digital twins	4	5	3
Flexible manufacturing systems	5	4	3
AR/VR support at work	3	4	4
Autonomous mobile robots	4	4	3

The results are consistent with prior research, highlighting high technical and cognitive demands in smart manufacturing (Thürer et al., 2022), the importance of psychosocial competencies in Industry 4.0 transitions

(Ergün and Cevickan, 2021), and the growing role of human adaptability in Industry 5.0 environments (Romero et al., 2021). The table is based on the analysis of technological complexity, competency clusters, relevant literature, and expert judgment, and represents a preliminary framework for further validation.

4 Theoretical Prediction of Competency Weights

In collaborative robotics, understanding the competencies of different user groups is essential for the successful implementation of new technologies. Before conducting the empirical study, it is important to predict the relative importance of competencies for each target group (experts, students, novices). Using standardized frameworks such as ISO 22156 (International Organization for Standardization, 2022), ISO 9241-210 (International Organization for Standardization, 2019), and ISO/TS 15066 (International Organization for Standardization, 2016), together with AHP and TOPSIS methods, enables the definition of competency significance in advance. Predefined weights improve methodological transparency and support empirical verification of competency assumptions across target groups (MacCoun, 2003). They also provide a basis for designing tasks and assessment tools aligned with the expected importance of competencies, supporting the development of suitable profiles for working with collaborative robots and identifying groups that require additional training or support.

4.1 Prediction of weights using the ISO analytical model

Using the ISO analytical model, we derived the weights of technical, cognitive, and psychosocial competencies through a methodological review and analysis of ISO standards (Romero et al., 2022). The weights were determined through the review of key standards, qualitative categorization and quantification of content, averaging of results across standards, and adjustment according to professional profiles and practical recommendations (Buhl et al., 2021). The identification of competencies was based on the relevance of each standard to human–robot interaction and assessed on a scale from 0 to 1 according to ISO/IEC 24773−1 (International Organization for Standardization & International Electrotechnical Commission, 2022). Each selected ISO standard was analyzed in terms of explicit and implicit requirements for the individual competency categories. The ratios were determined based on:

4.1.1 Qualitative Content Analysis:

- ISO 22156 (technical: 0.6, cognitive: 0.3, psychosocial: 0.1) focuses primarily on technical areas such as understanding robot operations, programming, and integration. The emphasis on technical competencies is rated at 0.6.
- ISO 9241-210 (cognitive: 0.6, psychosocial: 0.3, technical: 0.1) prioritizes user experience, perception, cognitive load, and decision-making processes, which aligns with the cognitive competency rating of 0.6.
- ISO/TS 15066 (psychosocial: 0.5, cognitive: 0.3, technical: 0.2) addresses human limits, risk, hazard perception, and teamwork, justifying the emphasis on psychosocial competencies with a rating of 0.5.

4.1.2 Consideration of the proportion of text related to each category:

For each standard, we analyzed the proportion of content (e.g., number of paragraphs, chapters, or sections) that directly or indirectly addresses a specific competency. Based on this, a weight was assigned on a scale from 0 to 1. For clarity, Table 2 presents a comparative analysis of the content distribution within each standard. For example, ISO 22156:2021 addresses technical aspects of human-robot collaboration (operating modes, safety systems, robot structure, etc.) in more than half of its content, which justifies assigning a weight of 0.6 to technical competencies. In contrast, ISO 9241-210:2019 focuses on the design of interactive systems from the user's perspective, frequently referencing cognitive load, decision-making, and information perception, which supports a weight of 0.6 for cognitive competencies. This approach ensures traceability and reproducibility of the assessment while providing a methodologically grounded basis for subjective judgment. Naturally, the credibility of the model could be further reinforced using a modified Delphi method as an alternative or enhancement.

Table 2: Comparative Analysis of Content Distribution in Each Standard

Standard	Percentage of Content on Technical Competencies	Percentage of Content on Cognitive Competencies	Percentage of Content on Psychosocial Competencies
ISO 22156:2021	60	30	10
ISO 9241−210:2019	10	60	30
ISO/TS 15066:2016	20	30	50

In the next step, based on the content analysis of the mentioned ISO standards, we determined the base weights of the three competency categories as the average emphasis across the individual sources: technical competencies: 0.30, cognitive competencies: 0.40, psychosocial competencies: 0.30.

4.1.3 Consistency with Other Normative and Applied Sources

The initial competency values did not fully reflect differences in prior knowledge and participant roles; therefore, they were additionally balanced according to real industrial requirements, where technical competencies often dominate (Buhl et al., 2021). This adjustment was supported by recommendations from the Fraunhofer Institute for Manufacturing Engineering and Automation and the ARENA2036 project, which combine ISO-based normative frameworks with advanced industrial practices in human–robot collaboration (Romero et al., 2022; ISO/IEC 24773−1:2022). In addition to ISO 10218−1, ISO 10218−2, and ISO/TS 15066, the balancing process considered ISO 12100 (International Organization for Standardization, 2010), ISO 22935 (International Organization for Standardization, 2009), ISO 45001 (International Organization for Standardization, 2018), and ISO/IEC 24773−1 (International Organization for Standardization, 2022), covering risk assessment, occupational safety, psychosocial factors, and operator competencies. Competency weights were then adjusted according to participant roles and industrial requirements: experts emphasized technical competencies, students a balanced combination of technical and cognitive competencies, and novices mainly safety and psychosocial factors.

Table 3: Balanced Weight Values

Group	Technical Competencies	Cognitive Competencies	Psychosocial Competencies
Experts	0.50	0.35	0.15
Students	0.30	0.40	0.30
Laypeople	0.20	0.35	0.45

Fraunhofer IPA (2020) emphasizes that balanced weights better reflect the characteristics of different target groups by considering their knowledge, robotics experience, and task requirements. Such weighting supports the development of targeted training programs and competency-based learning paths. Romero et al. (2020) further highlight those personalized approaches in digitalized manufacturing environments can improve motivation and reduce resistance to technological change.

4.2 Prediction of competency weights using the AHP method

For systematic prediction of weights, Multi-Criteria Decision Making (MCDM) methods can also be used, representing a structured approach for decision-making in situations where multiple, often conflicting criteria must be considered. One of the most widely used MCDM methods is the Analytic Hierarchy Process (AHP). The results are organized into a comparison matrix $\boldsymbol{A}$, where element a_{ij} represents the relative importance of criterion *i* compared to criterion *j*. The relative importance of criteria *i* and *j* is assessed using Saaty's scale, with values ranging from 1 to 9. The comparison matrix is reciprocally symmetric, meaning $a_{ij} = 1/a_{ji}$. Matrices comparing the importance of individual competencies for each target group are constructed based on expert evaluations and guidelines (Romero et al., 2020; Buhl et al., 2021; Liu et al., 2022; Huang & Wang, 2023). The criteria used for pairwise comparisons are technical (T), cognitive (C), and psychosocial (P) competencies. Prediction of Competency Weights for Experts:

	T	C	P
T	1	2	3
C	0.5	1	2
P	0.33	0.5	1

Working with collaborative robots requires a specific set of technical skills, as experts are expected to interact directly with the robot and actively contribute to the optimization of production processes. Prediction of Competency Weights for Students:

	T	C	P
T	1	0.5	2
C	2	1	2
P	0.5	0.5	1

In working with collaborative robots, expert opinion places the greatest emphasis on cognitive competencies, which enable students to effectively apply their understanding of technology, adapt to changing conditions, and solve complex problems in various situations. Prediction of Competency Weights for Laypersons:

	T	C	P
T	1	0.5	0.25
C	2	1	0.5
P	4	2	1

Expert opinion places the greatest emphasis on psychosocial competencies (e.g., communication, emotional stability) and considerably less on technical competencies, as laypersons lack technical knowledge and cannot actively operate in the world of collaborative robotics.

Table 4: AHP Weight Value

Group	Technical Competencies	Cognitive Competencies	Psychosocial Competencies
Experts	0.54 (0.50)	0.30 (0.35)	0.16 (0.15)
Students	0.31 (0.30)	0.49 (0.40)	0.20 (0.30)
Laypersons	0.14 (0.20)	0.29 (0.35)	0.57 (0.45)

Table 4 compares competency weights obtained with the AHP method and balanced ISO model values. Experts show the highest technical competency weights in both models, students are dominated by cognitive competencies, and laypersons by psychosocial competencies. Results from AHP and ISO-based weights show similar trends across all groups, confirming the stability of competency assessments.

4.3 TOPSIS procedure with AHP-derived weights

The combination of the AHP and TOPSIS (Technique for Order Preference by Similarity to Ideal Solution) methods is considered one of the most effective approaches for multi-criteria decision-making (Opricovic & Tzeng, 2004). The AHP method determines criterion weights through pairwise comparisons, while TOPSIS identifies the optimal alternative according to its proximity to the ideal solution and distance from the worst solution. Their integration increases objectivity by combining systematic weight determination with quantitative ranking of alternatives (Sivalingam & Subramaniam, 2024; Tran et al., 2024). The procedure begins with defining the hierarchical structure of objectives, criteria, and alternatives. After determining weights with AHP, the TOPSIS method is applied through the construction of a decision matrix with normalized criteria values and assigned weights. The TOPSIS procedure then evaluates alternatives according to their distance from the ideal and worst solutions (Behzadian et al., 2012). In this study, the decision matrix of initial competency assessments (Table 5) contains hypothetical values based on expected group competencies and research assumptions from the literature (Fraunhofer IPA, 2020; ISO/TS 15066).

Table 5: Decision Matrix of Initial Competency Assessments

Group	Technical Competencies	Cognitive Competencies	Psychosocial Competencies
Experts	90	80	70
Students	70	75	65
Laypersons	50	60	80

Experts showed the highest technical competencies (90), followed by cognitive (80) and psychosocial competencies (70). Students relied most on cognitive competencies (75), while technical (70) and psychosocial competencies (65) were less emphasized because of limited practical experience. For laypersons, psychosocial competencies (80) were the most important, followed by cognitive (60), whereas technical competencies (50) represented the weakest area. The decision matrix was then normalized into a dimensionless form to ensure comparability between criteria (Behzadian et al., 2012).

Group	T	K	P
Experts	0.723	0.64	0.562
Students	0.562	0.6	0.522
Laypersons	0.402	0.48	0.643

The weighted normalized matrix $\boldsymbol{V} = [v_{ij}]$ is obtained by applying the criteria weights, multiplying the normalized matrix by the weight vector (Behzadian et al., 2012).

Group	T	K	P
Experts	0.2386	0.2304	0.1742
Students	0.1855	0.2160	0.1618
Laypersons	0.1326	0.1728	0.1993

Next comes the ranking step, aiming to identify the best alternative that satisfies all criteria, being closest to the ideal solution and farthest from the worst solution. The ideal solution A^+ includes the most favorable values for all criteria (higher values are better), while the worst solution A^- represents the least favorable values (lower values are worse) (Behzadian et al., 2012). Based on the weighted normalized matrix, the ideal and worst solutions are defined as follows: $A^+ = (0.2386,\ 0.2304,\ 0.1993)$ $A^- = (0.1326,\ 0.1728,\ 0.1618)$. The distance of each alternative from the ideal and worst solutions is then calculated, followed by the relative closeness index C_i, which enables the final ranking of alternatives (Behzadian et al., 2012). The results for experts, students, and laypersons are summarized in Table 6.

Table 6: The results of the calculations for experts, students, and laypersons are summarized

Group	S_i^+	S_i^-	C_i
Experts	0.0251	0.1212	0.828
Students	0.0644	0.0692	0.518
Laypersons	0.1170	0.0374	0.242

Experts achieved the highest score, confirming that strong technical and cognitive competencies provide the closest match to the ideal profile. Students ranked in the middle because of their theoretical foundation but limited practical experience, while laypersons were furthest from the ideal solution, indicating that psychosocial competencies alone cannot compensate for lower technical and cognitive abilities. Overall, the TOPSIS ranking confirms the validity of the AHP analysis and highlights the importance of balanced competency profiles in human–robot collaboration (Behzadian et al., 2012). Only minor differences were observed between AHP-derived and ISO-based weights, confirming the robustness of the model. Collaborative robots achieved the highest scores, followed by digital twins and autonomous mobile robots because of their technical and cognitive demands, whereas AR/VR systems and exoskeletons achieved lower scores due to their stronger psychosocial emphasis.

Table 7: TOPSIS Scores of Experts in Different Technologies

Technology	TOPSIS Score (AHP Weights)	TOPSIS Score (ISO Weights)
Collaborative Robots	1.0000	1.0000
Smart Exoskeletons	0.2292	0.2282
Digital Twins	0.4899	0.5038
Flexible Manufacturing Systems	0.7130	0.6887
AR/VR Work Support	0.1282	0.1276
Autonomous Mobile Robots	0.4379	0.4259

5 Conclusion

In this paper, we presented a theoretical framework to assess the key competencies required to work with collaborative robots. By analyzing job requirements and operators' tasks, we identified those with the greatest impact on safety, efficiency, and work quality. These tasks were then linked to the corresponding competencies, followed by a hierarchical ranking of competencies according to their importance. Using the multi-criteria decision-making methods AHP and TOPSIS, we systematically developed well-founded predictions that enable more targeted design of educational programs and the development of assessment tools tailored to the real needs of the work environment.

5.1 Results and discussion

The results of the study offer a structured approach to the development of a competency model and support decision-making in the design of training programs for work in modern smart factories. Theoretical predictions of competency weights contribute to a more deliberate design of practical tests used to assess the abilities of

candidates or employees. If certain competencies are identified in advance as more critical, the difficulty of tasks, required precision, and scope of responsibility can be adjusted accordingly. This is particularly relevant when comparing different target groups, such as experienced professionals, students with only theoretical knowledge, or laypersons without prior experience with robots. Such adjustments ensure greater relevance in the assessment of competency potential. Despite its many advantages, the proposed combination is not without limitations. Moreover, the model assumes static decision-making conditions, while in practice, the competency profiles of employees and technological requirements can change rapidly. Therefore, it would be reasonable for further research to focus on empirical validation of the model, for instance, by comparing its outcomes with the actual performance of individuals working with collaborative robots in real-world environments. Additionally, it would be useful to include a broader group of experts and conduct further consultations using the Delphi method to re-examine and validate the appropriateness of the selected criteria and their weights. A potential extension of the model involves the application of advanced methods such as fuzzy AHP or fuzzy TOPSIS, which allow for better handling of uncertainty and vagueness inherent in the subjective evaluation of competencies. Furthermore, future research could benefit from a comparative analysis with other multi-criteria decision-making methods, such as VIKOR, ELECTRE, or PROMETHEE, in order to test the robustness and broader applicability of the developed model across various contexts of industrial practice.

6 Ethics Declaration

The Commission for Research Ethics at Faculty of Economics and Business at University of Maribor establishes that the research titled Determining competencies for working with collaborative robots, proposed by Tadej Crnjac, falls under research with no risk to participants and issues a positive opinion regarding the ethical aspects of the research..

7 AI Declaration

During the preparation of this work, the authors used Grammarly and DeepL in order to polish and refine the text and grammar. After using these tools, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

References

Bahrin, M. A. K., Othman, M. F., Azli, N. N., Talib, M. F. (2016). Industry 4.0: A review on industrial automation and robotics. Journal of Industrial and Information Technology, 6(3), 34–39.

Bughin, J., Hazan, E., Ramaswamy, S., Chui, M., Allas, T., Dahlström, P., Trench, M. (2018). Skill shift: Automation and the future of the workforce. McKinsey Global Institute. Retrieved from https://www.mckinsey.com/featured-insights/future-of-work/skill-shift-automation-and-the-future-of-the-workforce

Buhl, H. U., Gabrek, N., Gerdes, J.-N., Kaymakci, C., Rauland, K., Richter, F., Sauer, A., Schneider, C., Schott, P., Seifermann, S., Tristán, A., Wagner, J., Wagon, F., Weibelzahl, M., Weissflog, J., & Zachmann, B. (2021). Industrial fexibility options and their applications in a future energy system. Fraunhofer FIT. https://doi.org/10.24406/fit-n-639062

Dabić, M., Stojčić, N., Simić, M., Potocan, V., Slavković, M., & Nedelko, Z. (2021). Intellectual agility and innovation in micro and small businesses: The mediating role of entrepreneurial leadership. *Journal of Business Research*, *123*, 683–695. https://doi.org/10.1016/j.jbusres.2020.10.013

Ergün, E., Cevikcan, E. (2021). *Competency modeling for Industry 4.0 transformation: A Delphi study and its application.* Computers & Industrial Engineering, 159, 107525.

Frank, A. G., Dalenogare, L. S., Ayala, N. F. (2019). Industry 4.0 technologies: Implementation patterns in manufacturing companies. International Journal of Production Economics, 210, 15–26. https://doi.org/10.1016/j.ijpe.2019.01.004

Fraunhofer IPA. (2020). *Human-Robot Collaboration – Best Practices and Recommendations for Industrial Applications*. Stuttgart: Fraunhofer Institute for Manufacturing Engineering and Automation IPA.

Frey, C. B., Osborne, M. A. (2017). The future of employment: How susceptible are jobs to computerisation? Technological Forecasting and Social Change, 114, 254–280. https://doi.org/10.1016/j.techfore.2016.08.019

Huang, Y., & Wang, J. (2023). Enhancing student competence in robotics: A cognitive perspective. Education and Information Technologies, 28(2), 867–890. https://doi.org/10.1007/s10639-022-10987-5

International Organization for Standardization. (2010). *ISO 12100:2010 Safety of machinery – General principles for design – Risk assessment and risk reduction*. https://www.iso.org/standard/51528.html

International Organization for Standardization. (2016). *ISO/TS 15066:2016 Robots and robotic devices – Collaborative robots*. https://www.iso.org/standard/62996.html

International Organization for Standardization. (2019). *ISO 9241-210:2019 Ergonomics of human-system interaction – Human-centred design for interactive systems*. https://www.iso.org/standard/77520.html

International Organization for Standardization & International Electrotechnical Commission. (2022). *ISO/IEC 24773-1:2022 Software engineering – Certification of software engineering professionals – Part 1: General requirements*. https://www.iso.org/standard/77614.html

Le Blanc, P. M., González-Romá, V., Wang, H. (2020). Stress in the workplace: Sources and implications for employee well-being. Journal of Occupational Health Psychol ogy, 25(3),183−198. https://doi.org/10.1037/ocp0000158

Liu, Y., et al. (2022). Adaptive learning in manufacturing: Trends and future perspectives. Journal of Manufacturing Systems, 62, 395–410. https://doi.org/10.1016/j.jmsy.2021.11.015

Liu, L., Guo, F., Zou, Z., Duffy, V. G. (2024). Application, Development and Future Opportunities of Collaborative Robots (Cobots) in Manufacturing: A Literature Review. International Journal of Human–Computer Interaction, 40(4), 915−932. https://doi.org/10.1080/10447318.2022.2041907

MacCoun, R. (2003). Experimental and quasi-experimental designs for generalized causal inference, by William R. Shadish, Thomas D. Cook, and Donald T. Campbell. *Journal of Policy Analysis and Management*, 22(2), 330–332. https://doi.org/10.1002/pam.10129

Ojsteršek, R., Buchmeister, B., & Javernik, A. (2024). Personalizing Human–Robot Workplace Parameters in Human-Centered Manufacturing. Machines, 12(8), 546. https://doi.org/10.3390/machines12080546

Opricovic, S., & Tzeng, G. H. (2004). Compromise solution by MCDM methods: A comparative analysis of VIKOR and TOPSIS. *European Journal of Operational Research*, *156*(2), 445–455. https://doi.org/10.1016/S0377-2217(03)00020-1

Potocan, V., Lazányi, K., & Nedelko, Z. (2024). Industry 4.0 and human–robot interaction (pp. 87−100). https://doi.org/10.4324/9781003221050-6

Romero, D., et al. (2021). *Towards a human-centric Industry 5.0: A human digital twin approach.* IFAC-PapersOnLine, 54(1), 607–612

Roychoudhury, A., Khorshidi, S., Agrawal, S., Bennewitz, M. (2023). Perception for Humanoid Robots. Current Robotics Reports, 4(4), 127−140. https://doi.org/10.1007/s43154-023-00107-x

Salas, E., Rosen, M. A. (2020). Building high-performance teams in human-robot collaboration. Journal of Applied Psychology, 105(3), 237–246. https://doi.org/10.1037/apl0000416

Segura, P., Lobato-Calleros, O., Ramírez-Serrano, A., & Soria, I. (2021). Human-robot collaborative systems: Structural components for current manufacturing applications. Advances in Industrial and Manufacturing Engineering, 3, 100060. https://doi.org/10.1016/J.AIME.2021.100060

Siciliano, B., Khatib, O. (2016). Robotics and the Handbook. In Springer Handbook of Robotics (pp. 1–6). Springer.

Sivalingam, S., & Subramaniam, S. (2024). Cobot selection using hybrid AHP-TOPSIS based multi-criteria decision making technique for fuel filter assembly process. *Heliyon*, *10*(3), e27145. https://doi.org/10.1016/j.heliyon.2024.e27145

Villani, V., Pini, F., Leali, F., & Secchi, C. (2018). Survey on human–robot collaboration in industrial settings: Safety, intuitive interfaces, and applications. Mechatronics, 55, 248−266. Dostopno na https://doi.org/10.1016/j.mechatronics.2018.02.009.

Wang, Q., Liu, H., Ore, F., Wang, L., Hauge, J. B., & Meijer, S. (2023). Multi-actor perspectives on human robotic collaboration implementation in the heavy automotive manufacturing industry: A Swedish case study. *Technology in Society*, *72*, 102165. https://doi.org/10.1016/j.techsoc.2022.102165

Welfare, K. S., Hallowell, M. R., Shah, J. A., Riek, L. D. (2019). Consider the Human Work Experience When Integrating Robotics in the Workplace. 2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI), 75−84. https://doi.org/10.1109/HRI.2019.8673139

Werning, E., Negri, M., Verweij, S., Rossi, G., Rataj, S., & Likar, B. (2021). MODEL SMEART in smernice za poslovno sodelovanje med univerzo in gospodarstvom. https://doi.org/10.26493/978-961-293-051-6

Wilson, M. A., & Salas, E. (2010). Task analysis techniques for human performance research. In S. G. Salas & D. M. Prieto (Eds.), Human performance and ergonomics: A critical review (pp. 121–136). CRC Press.

Unveiling the Unseen: A Meta-Synthesis of Qualitative Informal Female Entrepreneurship Research

Simona Mormile, Roberta Romano, Gabriella Piscopo, Paola Adinolfi
University of Salerno, Fisciano, Italy

Abstract

Entrepreneurship is widely recognized as a key driver of economic growth, innovation, and social development. However, traditional entrepreneurship research has largely focused on formal business activities, often overlooking the experiences of women operating within the informal economy. Informal female entrepreneurship (IFE) represents a significant yet underexplored phenomenon, particularly in developing contexts where structural barriers, limited employment opportunities, and social norms constrain women's participation in the formal labour market. Despite its growing relevance, scholarly attention to IFE remains fragmented, and qualitative insights into women's lived entrepreneurial experiences are still limited.

This study aims to advance the understanding of informal female entrepreneurship by conducting a meta-synthesis of qualitative research. Specifically, the study seeks to identify the key factors influencing women's engagement in informal entrepreneurial activities and to explore how these factors shape their entrepreneurial trajectories across different socio-economic contexts. To achieve this objective, a qualitative meta-synthesis methodology was adopted. Following a systematic search across major academic databases, 15 peer-reviewed qualitative studies published between 2011 and 2023 were selected according to predefined inclusion criteria. Data were extracted and analyzed using an open coding approach, allowing for the identification of recurring themes and conceptual categories across the selected studies.

The findings highlight four main dimensions shaping informal female entrepreneurship. First, women's access to labour market resources and entrepreneurial opportunities plays a crucial role in motivating engagement in informal ventures, often driven by limited access to formal employment or financial capital. Second, cultural and social barriers—such as gender norms, domestic responsibilities, and institutional discrimination—significantly influence women's entrepreneurial choices and growth potential. Third, digital technologies emerge as important enablers, providing new opportunities for networking, market access, and business development. Finally, informal female entrepreneurship generates significant economic and social contributions, including household income generation, community development, and women's empowerment.

By synthesizing qualitative evidence, this study contributes to the entrepreneurship literature in two main ways. First, it broadens the conceptualization of entrepreneurship by highlighting the diversity of women's entrepreneurial experiences within informal economic systems. Second, it provides a comprehensive overview of qualitative research on IFE, offering deeper insights into the motivations, strategies, and challenges faced by informal female entrepreneurs. The findings also suggest important implications for policymakers and practitioners, emphasizing the need for supportive policies, access to financial resources, and targeted training programs aimed at fostering women's entrepreneurial potential and facilitating pathways from informal to formal economic participation.

Keywords: Informal Female Entrepreneurship; Informal Sector; Female; Meta-synthesis; Qualitative Study.

Paper type: Academic Research Paper

1 Introduction

Entrepreneurship is widely recognized as a key driver of economic growth, innovation, and social development (Holcombe, 1998; Valliere and Peterson, 2009). Traditional research on entrepreneurship has largely focused on formal business activities characterized by legal registration, regulatory compliance, and access to financial markets (Baumol, 1993). However, an important part of entrepreneurial activity occurs outside formal institutional frameworks.

One emerging yet understudied phenomenon is informal female entrepreneurship (IFE), which refers to women engaging in entrepreneurial activities within the informal economy without formal registration, taxation, or regulatory compliance (Spring, 2009; Thapa Karki et al., 2021). Informal female entrepreneurship is particularly widespread in developing countries, where women often face limited access to formal employment opportunities and financial resources (Rashid and Ratten, 2020).

IFE encompasses a wide range of micro and small-scale activities, including street vending, home-based businesses, and small-scale service provision (Muhammad et al., 2021). These ventures frequently emerge as strategies for income generation and economic survival in contexts characterized by structural inequalities and restricted labor market participation. Despite operating outside formal regulatory frameworks, informal female entrepreneurs contribute significantly to household income, local economies, and community development (Brush and Cooper, 2012).

Several factors may encourage women to engage in informal entrepreneurship. Limited access to formal employment, cultural expectations, and gender-related barriers often push women toward informal economic activities (Agarwal and Lenka, 2015; Naguib, 2024). At the same time, informal entrepreneurship may offer women flexibility, autonomy, and opportunities for financial independence (Agyemang, 2017).

Despite the growing relevance of the informal sector in global economies, the literature on informal female entrepreneurship remains fragmented and relatively limited (Rugina, 2019). Existing studies often focus on the broader informal economy without specifically examining the gendered dynamics shaping women's entrepreneurial experiences. Furthermore, many studies rely on quantitative approaches, leaving a gap in qualitative research that explores the lived experiences, motivations, and challenges faced by women entrepreneurs operating informally (Corrêa et al., 2022).

Understanding the drivers and implications of IFE is particularly important given its potential impact on women's economic empowerment and socio-economic development. Informal entrepreneurship may serve as both a survival strategy and a pathway to opportunity, depending on contextual conditions and individual motivations (Amorós and Bosma, 2014).

In response to these gaps, this study conducts a meta-synthesis of qualitative research on informal female entrepreneurship. By synthesizing findings from qualitative studies, the research aims to identify recurring themes and factors shaping women's entrepreneurial activities in informal contexts.

Specifically, the study addresses the following research questions:

- What factors influence informal female entrepreneurship?
- How do these factors shape women's entrepreneurial experiences across different contexts?

Through this analysis, the paper seeks to contribute to the entrepreneurship literature by offering a more comprehensive understanding of women's participation in the informal economy and highlighting the broader socio-economic implications of informal female entrepreneurship.

The remainder of the paper is structured as follows. The next section reviews the relevant literature on informal female entrepreneurship. This is followed by the methodology describing the meta-synthesis approach. Subsequently, the results are presented and discussed, highlighting the main themes emerging from the analysis. Finally, the paper concludes with theoretical and practical implications for research and policy.

2 Literature Review

Informal female entrepreneurship has gained increasing attention in recent years due to its economic and social relevance. The phenomenon refers to women who establish and manage businesses within the informal economy, typically without official registration or compliance with formal regulatory frameworks (Williams and Nadin, 2010). These activities range from home-based production to small-scale trade and service provision.

Several studies highlight the significant role women entrepreneurs play in economic development. Female entrepreneurship contributes to job creation, innovation, and income generation, particularly in developing economies (Brush et al., 2019). Within the informal economy, women often engage in entrepreneurial activities as a strategy to overcome barriers in formal labour markets.

IFE is characterized by several key features. First, informal enterprises typically operate without formal registration or licensing (Agarwal and Lenka, 2015). Second, women entrepreneurs in the informal sector often face limited access to financial resources, training, and business networks (Minniti and Naudé, 2010). Despite these constraints, informal female entrepreneurs frequently demonstrate high levels of adaptability and resilience, relying on creativity and social networks to sustain their businesses (Jennings and Brush, 2013).

Another important dimension concerns the motivations driving women to engage in informal entrepreneurship. Previous research suggests that both necessity and opportunity factors may influence entrepreneurial entry. For some women, informal entrepreneurship represents a response to unemployment, economic necessity, or lack of formal employment opportunities. For others, it provides opportunities for autonomy, flexibility, and work-life balance (Williams, 2009; Amorós and Bosma, 2014).

A substantial body of literature also highlights the barriers faced by women entrepreneurs in informal contexts. These barriers include cultural norms, gender discrimination, limited access to financial capital, and difficulties in balancing entrepreneurial activities with domestic responsibilities (Ahl, 2006; Hughes et al., 2012). In many societies, traditional gender roles restrict women's economic participation and limit their ability to access entrepreneurial resources.

In addition to economic barriers, institutional factors may also influence informal female entrepreneurship. Institutional theory suggests that both formal regulations and informal social norms shape entrepreneurial behavior (Kazumi and Kawai, 2017). In contexts where formal institutions provide limited support for women entrepreneurs, informal economic activities may become an accessible alternative.

Despite these challenges, research highlights the significant contributions of informal female entrepreneurs to household welfare and community development. Women's entrepreneurial activities often play a crucial role in supporting family income and improving living conditions (Brush and Cooper, 2012). Moreover, entrepreneurial engagement may enhance women's empowerment and social status within their communities (Amorós and Bosma, 2014).

In recent years, scholars have also begun to explore the role of policy interventions and support mechanisms in promoting women's entrepreneurship. Programs such as microfinance initiatives, entrepreneurial training, and regulatory reforms have been proposed to facilitate the growth and formalization of women-led businesses (Datta and Gailey, 2012).

However, despite the growing interest in the topic, several gaps remain in the literature. In particular, qualitative studies on informal female entrepreneurship remain relatively limited and fragmented. Qualitative approaches are particularly valuable for understanding the motivations, strategies, and lived experiences of women operating in informal contexts (Corrêa et al., 2022). By capturing context-specific dynamics and individual perspectives, qualitative research can provide deeper insights into the processes underlying informal entrepreneurship.

To address these gaps, the present study adopts a meta-synthesis approach to integrate findings from existing qualitative studies. This approach allows for the identification of common themes and patterns across different contexts, contributing to a more comprehensive understanding of informal female entrepreneurship.

3 Methodology

To explore the dynamics of informal female entrepreneurship (IFE), this study adopts a qualitative meta-synthesis approach, which allows the systematic integration and interpretation of findings from multiple qualitative studies (Hoon, 2013; Walsh and Downe, 2005). Unlike meta-analysis, which statistically aggregates quantitative data, meta-synthesis focuses on interpreting qualitative findings to identify common themes and theoretical insights across studies.

Our methodological approach was based on five sequential steps, represented graphically in Figure 1, and explained in detail in the following subsections, involving synthesizing qualitative arguments to create a model.

The objective was to explore the key factors influencing informal female entrepreneurship and to understand how these factors shape women's entrepreneurial experiences across different contexts.

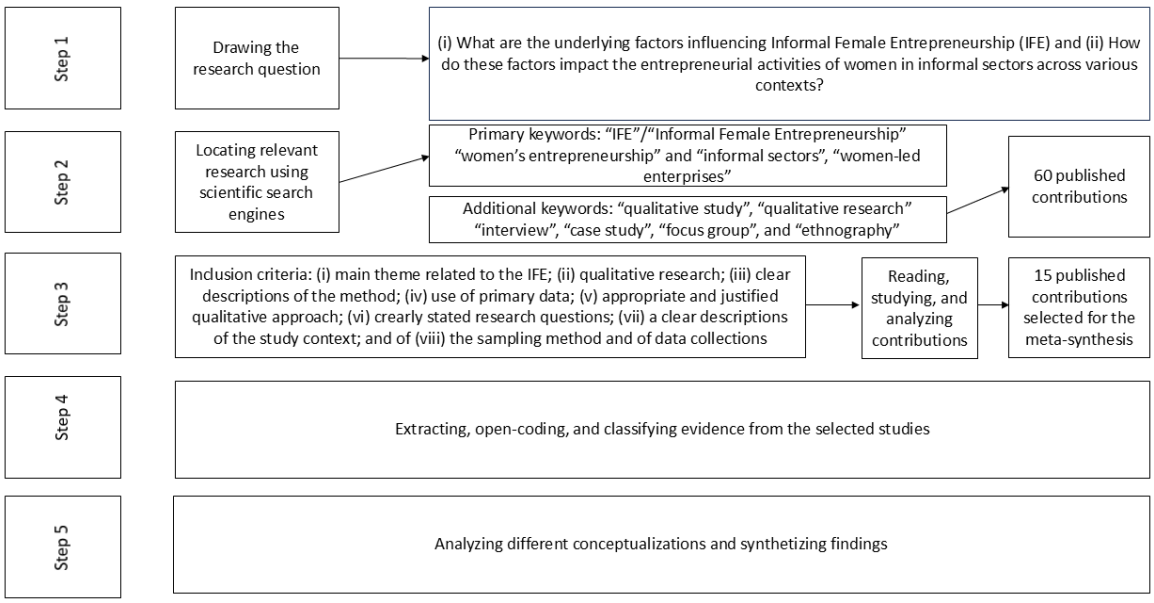


Figure 1 The five methodological steps of meta-synthesis

3.1 Research Design and Research Questions

The analysis was guided by two research questions:

- What factors influence informal female entrepreneurship?
- How do these factors shape women's entrepreneurial activities in informal sectors across different contexts?

These questions guided the identification and analysis of qualitative studies examining women's experiences in informal entrepreneurial activities.

3.2 Study Selection

Relevant studies were identified through a systematic search of major academic databases, including Scopus, Web of Science, EBSCOHost Business Source Complete, and Google Scholar. The search included keywords such as *informal female entrepreneurship*, *women's entrepreneurship*, *informal sector*, and *women-led enterprises*, combined with terms related to qualitative research (e.g., *interview*, *case study*, *focus group*, and *ethnography*).

The search was conducted between October 2023 and January 2024 and was complemented by manual searches of reference lists and relevant academic networks.

To ensure the relevance and quality of the selected studies, the following inclusion criteria were applied:

- focus on informal female entrepreneurship;
- use of a qualitative research design;
- primary data collection (e.g., interviews, observations, or focus groups).
- clear description of research context and methodology

The initial search identified 60 studies. After screening titles and abstracts and reviewing full texts, 15 qualitative studies published between 2011 and 2023 met the inclusion criteria and were included in the meta-synthesis (Table 1).

Table 1 Details of studies selected for meta-synthesis

Authors	**Year of Publication**	**Title**	***Journal/Book***	**Doi**
Ahsan, R.	2023	Female Academic Entrepreneurship: Evidence From a Developing Country	*New Horizons and Global Perspectives in Female Entrepreneurship Research*	10.1108/978-1-83982-780-820231005
Franck, A. K.	2012	Factors motivating women's informal micro-entrepreneurship: Experiences from Penang, Malaysia.	*International Journal of Gender and Entrepreneurship*	10.1108/17566261211202981
Kalafatoglu, T., Mendoza, X.	2017	The impact of gender and culture on networking and venture creation: An exploratory study in Turkey and MENA region	*Cross Cultural & Strategic Management*	10.1108/CCSM-04-2016-0090
Karakire Guma, P.	2015	Business in the urban informal economy: Barriers to women's entrepreneurship in Uganda	*Journal of African Business*	10.1080/15228916.2015.1081025
King, A.	2016	Access to opportunity: A case study of street food vendors in Ghana's urban informal economy	*Gender and Food: From Production to Consumption and After*	10.1108/S1529-212620160000022014
Korzenevica, M., Grasham, C. F., Johnson, Z., Gebreegzabher, A., Mebrahtu, S., Zerihun, Z., Charles, K. J.	2022	Negotiating spaces of marginality and independence: On women entrepreneurs within Ethiopian urbanization and water precarity	*World Development*	10.1016/j.worlddev.2022.105966
Kwami, J. D.	2015	Gender, entrepreneurship, and informal markets in Africa: Understanding how Ghanaian women traders self-organize with digital tools	*Economics: Concepts, Methodologies, Tools, and Applications*	10.4018/978-1-4666-8468-3.ch042

Lock, R., Lawton Smith, H.	2016	The impact of female entrepreneurship on economic growth in Kenya	*International Journal of Gender and Entrepreneurship*	10.1108/IJGE-11-2015-0040
Mehtap, S., Ozmenekse, L., Caputo, A.	2019	"I'm a stay at home businesswoman": an insight into informal entrepreneurship in Jordan	*Journal of Entrepreneurship in Emerging Economies*	10.1108/JEEE-10-2017-0080
Mukherjee, S., Pathak, S.	2023	Promotion of Women Micro Entrepreneurship and the Role of NPDAs: Case Study From India	*New Horizons and Global Perspectives in Female Entrepreneurship Research*	10.1108/978-1-83982-780-820231006
Mukit, M. M. H., Nabila, N. J., Islam, M. S., Abdel-Razzaq, A. I.	2020	Valuing women entrepreneurship in the informal economy in Bangladesh	*Economics-innovative and economics research journal*	10.2478/eoik-2020-0020
Quak, E. Barenboim, I.	2022	Female Entrepreneurship and Informality in Low- and Middle-Income Countries: What Have We Learned So Far?	*Institute of Development Studies*	10.19088/MUVA.2022.001
Shava, H., Chinyamurindi, W. T.	2022	Barriers of growth within an informal sector business: narratives of women subsistence entrepreneurs in South Africa	*African Journal of Economic and Management Studies*	10.1108/AJEMS-08-2021-0354
Williams, C. C., Gurtoo, A.	2011	Evaluating women entrepreneurs in the informal sector: Some evidence from India	*Journal of Developmental Entrepreneurship*	10.2139/ssrn.2290543
Yusuff, O. S.	2013	The dynamics of strategic entry and motivations of yoruba female textile traders in the Balogun Market, Lagos State, Nigeria	*Journal of Developmental Entrepreneurship*	10.1142/S108494671350012X

3.3 Data Analysis

The selected studies were analyzed using an open coding approach to identify key themes and recurring concepts across the literature (Strauss and Corbin, 1998). The analysis focused primarily on the findings and discussion sections of the selected studies in order to capture the main insights generated by the original researchers.

Through an iterative analytical process, the identified codes were grouped into broader conceptual categories that reflected recurring patterns across the studies. This approach allowed the analysis to move from individual study findings to a cross-study interpretation, generating higher-level theoretical insights.

The final step of the analysis involved synthesizing the emerging categories into broader thematic dimensions that explain the main factors shaping informal female entrepreneurship. These dimensions form the basis of the findings presented in the following section.

4 Results

The meta-synthesis of the selected qualitative studies identifies several recurring themes that explain the dynamics of informal female entrepreneurship (IFE). The findings highlight four main dimensions: access to labour market resources, cultural and social barriers, the role of digital technologies, and the economic and social contributions of informal female entrepreneurs.

4.1 Access to Labor Market Resources and Opportunities

Female entrepreneurship represents an important pathway toward economic independence and empowerment in many contexts. Several studies highlight that women engage in informal entrepreneurship as a response to limited access to formal employment and financial resources.

Research on informal entrepreneurs in Ghana and Ethiopia shows how women demonstrate strong resilience and creativity in overcoming structural constraints. In many cases, women intentionally pursue informal entrepreneurship as a strategic choice to improve their economic conditions rather than merely as a survival strategy.

Education and professional training also play a crucial role in enabling women to develop entrepreneurial skills and expand their opportunities. However, access to financial capital remains a persistent challenge. Studies conducted in Uganda and India show that women entrepreneurs frequently encounter difficulties obtaining credit and financial support due to institutional barriers and gender biases.

In addition to financial constraints, limited access to markets, networks, and mentorship opportunities further restricts the growth potential of women-led informal businesses. These findings highlight the importance of developing more inclusive support systems and entrepreneurial infrastructures.

4.2 Cultural and Social Barriers

Cultural norms and traditional gender roles represent significant barriers to women's entrepreneurial activities. In many contexts, social expectations regarding women's domestic responsibilities limit their participation in economic activities and restrict their access to entrepreneurial resources.

Studies conducted in the MENA region and Sub-Saharan Africa demonstrate how gender norms influence women's entrepreneurial opportunities and constrain their business development. Women entrepreneurs often face discrimination, limited access to financial services, and additional pressures related to family responsibilities.

These structural barriers are frequently reinforced by institutional and societal dynamics that limit women's participation in economic life. Addressing these challenges requires not only economic support measures but also broader cultural and social interventions aimed at promoting gender equality and changing societal perceptions regarding women's entrepreneurship.

4.3 Digital Technologies as Enablers of Informal Female Entrepreneurship

Digital technologies increasingly play an important role in supporting women entrepreneurs operating in informal sectors. Online platforms and digital tools allow women to expand their market access, develop professional networks, and improve business operations.

Evidence from studies conducted in Jordan and Ghana highlights how digital technologies can help women overcome physical and social barriers that limit traditional business opportunities. Through digital platforms, entrepreneurs can access new customers, acquire information, and develop innovative business strategies.

These findings suggest that digital literacy and technological access represent important factors in supporting the growth and sustainability of informal female entrepreneurship.

4.4 Economic and Social Contributions

Despite operating outside formal regulatory frameworks, informal female entrepreneurs contribute significantly to economic and social development. Informal enterprises generate income for households, create employment opportunities, and support local economic systems.

Studies conducted in India, Bangladesh, and Nigeria demonstrate that women's informal entrepreneurial activities often contribute to family welfare, community development, and social cohesion. In addition to their economic contributions, women entrepreneurs frequently play important roles in supporting family education, community networks, and cultural preservation.

Overall, the findings highlight the importance of recognizing informal female entrepreneurship as a significant component of economic activity and a potential driver of women's empowerment and social inclusion.

5 Discussion and Conclusions

The purpose of this study was to explore the dynamics of informal female entrepreneurship through a meta-synthesis of qualitative research. By analyzing fifteen qualitative studies, this research identifies key factors shaping women's entrepreneurial activities within informal economic contexts.

The findings highlight the complex interplay between structural constraints and individual agency in shaping informal female entrepreneurship. Limited access to labor market opportunities and financial resources often encourages women to pursue informal entrepreneurial activities. At the same time, cultural norms and gender-related barriers continue to restrict women's access to entrepreneurial resources and opportunities.

Despite these challenges, the study highlights the resilience and adaptability of women entrepreneurs operating in informal contexts. Many women strategically leverage available resources, social networks, and digital technologies to develop sustainable entrepreneurial activities.

Digital technologies, in particular, emerge as important tools for overcoming traditional barriers to entrepreneurship. By facilitating access to information, markets, and professional networks, digital platforms provide new opportunities for women entrepreneurs to expand their businesses and enhance their economic participation.

In addition to their individual benefits, informal female entrepreneurs contribute significantly to broader socio-economic development. Their activities support household income, promote community welfare, and enhance women's empowerment in contexts where formal employment opportunities remain limited.

This study contributes to the literature on entrepreneurship in two main ways. First, it advances understanding of informal female entrepreneurship by synthesizing qualitative evidence on the factors shaping women's entrepreneurial activities in informal contexts. In doing so, it extends traditional entrepreneurship research, which has often focused primarily on formal business environments.

Second, the study highlights the value of qualitative research in capturing the lived experiences and motivations of women entrepreneurs. By integrating insights from multiple qualitative studies, this research provides a more comprehensive and nuanced understanding of informal female entrepreneurship.

The findings of this study have several implications for policymakers, educators, and practitioners. Policymakers should develop policies that recognize the unique characteristics of informal female entrepreneurship and facilitate access to financial resources, training programs, and entrepreneurial support services.

Educational institutions can also play a key role by promoting entrepreneurship education and digital skills development, particularly among women. Strengthening entrepreneurial competencies may help women transition from informal to formal economic activities.

Finally, women entrepreneurs themselves may benefit from expanding their use of digital technologies and professional networks to overcome traditional barriers and access broader market opportunities.

This study has several limitations. The meta-synthesis focuses exclusively on published qualitative studies, which may exclude relevant insights from grey literature and policy reports. Additionally, the studies included in the analysis are primarily concentrated in developing countries, potentially limiting the generalizability of the findings.

Future research could adopt mixed method approaches to combine qualitative insights with quantitative evidence, providing a more comprehensive understanding of informal female entrepreneurship. Longitudinal studies may also offer valuable insights into the evolution and sustainability of women-led informal businesses over time.

Further research should also explore the processes through which informal entrepreneurial activities transition into formal business ventures, as this transition represents an important pathway for economic development and women's empowerment.

6 Ethics declaration

The authors declare that ethical approval was not required for the research presented in this paper. The study does not involve human participants, personal data collection, or any procedures requiring ethical clearance according to institutional or national guidelines.

7 AI declaration

The authors declare that no artificial intelligence (AI) tools were used to generate or produce the content of this paper. All analyses, interpretations, and written material were developed independently by the authors.

References

Agarwal, S. and Lenka, U. (2015) *"Study on Work-Life Balance of Women Entrepreneurs – Review and Research Agenda",* Industrial and Commercial Training, Vol. 47, No. 7, pp. 356–362.

Agyemang, L. (2017) "Women's Work: Labor Market Outcomes and Female Entrepreneurship in Ghana", *Butler Journal of Undergraduate Research*, Vol. 3, No. 1.

Ahl, H. (2006) "Why Research on Women Entrepreneurs Needs New Directions", *Entrepreneurship Theory and Practice*, Vol. 30, No. 5, pp. 595–621.

Amorós, J.E. and Bosma, N. (2014) *Global Entrepreneurship Monitor 2013 Global Report: Fifteen Years of Assessing Entrepreneurship Across the Globe*, Global Entrepreneurship Research Association (GERA), Universidad del Desarrollo and Utrecht University.

Baumol, W.J. (1993) "Formal Entrepreneurship Theory in Economics: Existence and Bounds", *Journal of Business Venturing,* Vol. 8, No. 3, pp. 197–210.

Brush, C.G. and Cooper, S.Y. (2012) "Female Entrepreneurship and Economic Development: An International Perspective", *Entrepreneurship & Regional Development,* Vol. 24, No. 1, pp. 1–6.

Brush, C., Edelman, L., Manolova, T. and Welter, F. (2019) "A Gendered Look at Entrepreneurship Ecosystems", *Small Business Economics*, Vol. 53, pp. 393–408.

Corrêa, V.S., Brito, F.R.D.S., Lima, R.M.D. and Queiroz, M.M. (2022) "Female Entrepreneurship in Emerging and Developing Countries: A Systematic Literature Review", *International Journal of Gender and Entrepreneurship*, Vol. 14, No. 3, pp. 300–322.

Datta, P.B. and Gailey, R. (2012) "Empowering Women Through Social Entrepreneurship: Case Study of a Women's Cooperative in India", *Entrepreneurship Theory and Practice*, Vol. 36, No. 3, pp. 569–587.

Holcombe, R.G. (1998) "Entrepreneurship and Economic Growth", *Quarterly Journal of Austrian Economics*, Vol. 1, No. 2, pp. 45–62.

Hoon, C. (2013) "Meta-Synthesis of Qualitative Case Studies: An Approach to Theory Building", *Organizational Research Methods*, Vol. 16, No. 4, pp. 522–556.

Hughes, K.D., Jennings, J.E., Brush, C., Carter, S. and Welter, F. (2012) "Extending Women's Entrepreneurship Research in New Directions", *Entrepreneurship Theory and Practice*, Vol. 36, No. 3, pp. 429–442.

Jennings, J.E. and Brush, C.G. (2013) "Research on Women Entrepreneurs: Challenges to (and from) the Broader Entrepreneurship Literature?", *Academy of Management Annals*, Vol. 7, No. 1, pp. 663–715.

Kazumi, T. and Kawai, N. (2017) "Institutional Support and Women's Entrepreneurial Self-Efficacy*", Asia Pacific Journal of Innovation and Entrepreneurship*, Vol. 11, No. 3, pp. 345–365.

Minniti, M. and Naudé, W. (2010) "What Do We Know About the Patterns and Determinants of Female Entrepreneurship Across Countries?", *European Journal of Development Research*, Vol. 22, pp. 277–293.

Muhammad, S., Kong, X., Saqib, S.E. and Beutell, N.J. (2021) "Entrepreneurial Income and Wellbeing: Women's Informal Entrepreneurship in a Developing Context", *Sustainability*, Vol. 13, No. 18.

Naguib, R. (2024) "Motivations and Barriers to Female Entrepreneurship: Insights from Morocco", *Journal of African Business*, Vol. 25, No. 1, pp. 9–36.

Rashid, S. and Ratten, V. (2020) "Women Entrepreneurship in Emerging Economies: A Systematic Literature Review", in Managing Innovation and Creative Capabilities, Springer, pp. 37–88.

Rugina, S. (2019) "Female Entrepreneurship in the Baltics: Formal and Informal Context", *International Journal of Gender and Entrepreneurship*, Vol. 11, No. 1, pp. 58–74.

Spring, A. (2009) "African Women in the Entrepreneurial Landscape: Reconsidering the Formal and Informal Sectors*", Journal of African Business,* Vol. 10, No. 1, pp. 11–30.

Strauss, A. and Corbin, J. (1998) Basics of Qualitative Research Techniques, Sage Publications, London.

Thapa Karki, S., Xheneti, M. and Madden, A. (2021) "To Formalize or Not to Formalize: Women Entrepreneurs' Sensemaking of Business Registration in the Context of Nepal", *Journal of Business Ethics*, Vol. 173, pp. 687–708.

Valliere, D. and Peterson, R. (2009) "Entrepreneurship and Economic Growth: Evidence from Emerging and Developed Countries", *Entrepreneurship & Regional Development*, Vol. 21, No. 6, pp. 459–480.

Walsh, D. and Downe, S. (2005) "Meta-Synthesis Method for Qualitative Research: A Literature Review*", Journal of Advanced Nursing*, Vol. 50, No. 2, pp. 204–211.

Williams, C.C. (2009) "Evaluating the Extent and Nature of 'Envelope Wages' in the European Union: A Geographical Analysis", *European Spatial Research and Policy*, Vol. 16, No. 1, pp. 115–129.

Williams, C.C. and Gurtoo, A. (2011) "Evaluating Women Entrepreneurs in the Informal Sector: Some Evidence from India", *Journal of Developmental Entrepreneurship*, Vol. 16, No. 3, pp. 351–369.

Mapping Energy Innovation and ESG: A Bibliometric Review

Giovanni Spatola, Stefania D'Aprile, Antonio Capasso, Maria Zifaro, Andrea Presciutti
Mercatorum University, Rome, Italy

Abstract

The global energy sector is currently undergoing a profound "twin transition", characterized by the simultaneous pursuit of digital transformation and environmental sustainability. This dual trajectory is reshaping not only the technological infrastructure of energy production and distribution, but also the strategic, financial, and organizational logics governing utilities, independent producers, grid operators, and the wide ecosystem of stakeholders revolving around them. In this context, the integration between technological innovation and Environmental, Social, and Governance (ESG) considerations has emerged as a central axis around which contemporary academic and managerial debate revolves. The present study provides a comprehensive bibliometric analysis of the intersection between the energy sector, technological and organizational innovation, and performance metrics, specifically ESG factors, social acceptance, and profitability. The choice of a bibliometric approach is motivated by the need to overcome the fragmentation that historically characterizes this literature, where engineering, financial, managerial, and policy contributions tend to coexist without systematic integration. Using the Web of Science (WoS) Core Collection, a final dataset of 395 articles published between 2015 and 2026 was extracted, following a rigorous selection process inspired by the PRISMA protocol. The temporal window was deliberately set to coincide with the signing of the Paris Agreement and the adoption of the United Nations 2030 Agenda, two milestones that catalyzed academic attention and corporate strategic reorientation toward sustainability. Through the R package Bibliometrix and its graphical interface Biblioshiny, the research maps the intellectual structure, evolutionary trends, and collaboration networks underpinning the field, combining performance metrics, keyword co-occurrence networks, co-citation analysis, and historiographic mapping.

The results show an annual growth rate of 8.28% and an average of 26.31 citations per document, indicating a particularly dynamic and high-impact research field. The study identifies a clear thematic shift: from initial technical challenges to systemic themes (smart grids, regulatory policy), up to the recent emphasis on ESG metrics. Two main intellectual domains emerge: innovation systems for socio-technical transitions and sustainable business models based on the circular economy. Geographically, China leads in volume while Europe—particularly Italy and the United Kingdom—drives qualitative excellence, supported by Switzerland's disproportionately high citation impact. The paper concludes by outlining future research trajectories addressing the social dimension of ESG, Scope 3 emissions, and greenwashing detection.

Keywords: Circular Economy, Renewable Energy, ESG, Smart Grid, Bibliometrics.

Paper type: Academic Research Paper

1 Introduction

In recent decades, the global economic and industrial landscape has undergone radical transformations, but no sector has been subjected to as much pressure and profound change as the energy sector. Currently, the energy sector is at the center of what scholars and policymakers define as the "twin transition," a dual transition that combines digital transformation with the imperative need for environmental sustainability. This convergence represents not just a technological upgrade, but a true paradigm shift that redefines business models, value chains, and interactions with stakeholders. The implications cut across multiple dimensions: from the architecture of energy infrastructures to the design of regulatory frameworks, from capital allocation decisions to the renegotiation of the social contract between energy producers and the communities they serve.

The urgency to mitigate climate change, combined with the United Nations Sustainable Development Goals (SDGs) and international agreements like the Paris Agreement, has pushed energy companies, utilities, and governments to invest heavily in renewable energy, energy efficiency, and smart grids. However, technological innovation alone is not sufficient to ensure a successful transition. The need to integrate organizational and managerial innovation has strongly emerged, measuring results not only in terms of financial profit but also through holistic metrics. In this context, ESG (Environmental, Social, and Governance) criteria have become the gold standard for evaluating corporate sustainability performance, influencing access to capital, social acceptance, and institutional legitimacy. The financial system itself is being reshaped: green bonds, sustainability-linked loans, and ESG-screened investment funds have grown exponentially, channeling capital toward firms that can credibly demonstrate alignment with sustainability principles.

Despite growing academic and managerial interest in the intersection of energy innovation and ESG factors, the literature in this field appears fragmented. Studies tend to focus on disciplinary silos: some purely analyze the engineering aspects of smart grids or renewable energy, others focus on policy dynamics, while still others explore the financial impact of ESG metrics in isolation. There is a lack of a comprehensive overview that synthesizes how these research streams are merging and what the future trajectories are. Moreover, the rapid pace at which new contributions appear makes it increasingly difficult for both scholars and practitioners to navigate this body of knowledge without a structured cartographic tool.

To fill this gap, the present study aims to conduct a systematic and comprehensive bibliometric review. Bibliometrics, supported by powerful data analysis tools, allows for the deciphering of large volumes of scientific publications, identifying hidden networks, emerging themes, and the intellectual structure of a given research field. Unlike narrative reviews, bibliometric analyses minimize subjective biases and provide replicable, quantitatively grounded insights.

The Research Questions (RQ) guiding this study are as follows:

- *RQ1: What are the main publication trends, most prolific authors, leading countries, and highest-impact sources in the field of energy innovation and ESG?*
- *RQ2: How has the thematic structure of the research evolved over time, and what are the emerging, niche, and foundational themes?*
- *RQ3: What are the intellectual foundations (co-citation networks), and how are collaboration networks structured among researchers and institutions globally?*

By answering these questions, the paper offers a significant contribution to the literature, providing a conceptual map that will help future researchers position their studies and managers understand the evolutionary dynamics of the sector. Furthermore, the analysis is intended to serve policymakers seeking evidence-based orientation in the design of incentives, regulations, and standards capable of accelerating an authentic and inclusive energy transition.

2 Theoretical Background

To fully understand the results of the bibliometric analysis, it is necessary to outline the theoretical boundaries of the key concepts guiding this research: innovation in the energy sector, the circular economy, and the emerging role of ESG metrics. These three domains, while traditionally treated separately, have increasingly converged in both academic discourse and corporate practice, generating new hybrid frameworks that this study aims to map.

2.1 Innovation in the energy sector and socio-technical transitions

Innovation in the energy sector does not follow linear paths. It is best described through the lens of socio-technical transitions theory (Geels, 2002). According to this perspective, system changes (such as the shift from fossil fuels to renewable energy) occur through the interaction of three levels: "niches" (where radical innovations emerge), "regimes" (dominant practices, rules, and technologies), and the "landscape" (exogenous macro-trends like climate change or geopolitical crises). The interplay among these three levels explains why technological breakthroughs alone rarely translate into systemic change, and why successful transitions require coordinated alignment of technologies, institutions, user practices, and cultural meanings.

The introduction of technologies such as smart grids, energy storage systems, and distributed generation represents a direct challenge to traditional centralized energy regimes. Furthermore, the Technological Innovation Systems (TIS) approach (Bergek et al., 2008) emphasizes that the success of a new energy technology depends not only on its technical performance but also on the creation of actor networks, resource mobilization, guidance of the search, and institutional legitimation. Social acceptance, both at the local community level (e.g., for the installation of wind or biogas plants) and at the market level, thus becomes a critical success factor (Markard et al., 2016). Recent contributions have further enriched this framework by incorporating notions of justice, inclusion, and procedural legitimacy, recognizing that transitions inevitably generate winners and losers and that failure to address distributional impacts can stall or reverse otherwise promising trajectories.

2.2 Circular Economy and Sustainable Business Models

Parallel to technological innovation, the energy sector is experiencing profound innovation in business models, driven by the principles of the Circular Economy (CE). The circular economy aims to overcome the linear "take-make-dispose" model, aiming to close, slow, and narrow resource loops (Geissdoerfer et al., 2017). In the energy context, this translates into practices such as waste heat recovery, biomethane production from

agricultural waste (Cucchiella et al., 2018), recycling of solar panels and lithium batteries, and the integration of green supply chain management practices (Ghosh, 2022). These practices are no longer marginal experiments but increasingly form the backbone of competitive strategies in mature renewable industries.

From a strategic management perspective, the adoption of circular and sustainable business models can be interpreted through the Resource-Based View (RBV) (Barney, 1991). According to the RBV, companies that develop unique, rare, and inimitable capabilities in sustainable resource management and green innovation can achieve a sustainable competitive advantage. Sustainability ceases to be a mere regulatory compliance cost and becomes a driver of value creation. The dynamic capabilities perspective further extends this view, emphasizing the importance of sensing, seizing, and reconfiguring resources in response to rapidly evolving sustainability challenges.

2.3 The role of ESG metrics and performance measurement

The third theoretical pillar of this study concerns performance measurement through ESG criteria. While in the past the evaluation of energy companies was based almost exclusively on financial metrics (ROI, ROE, cash flows), today investors, regulators, and consumers demand transparency on environmental impacts (e.g., CO2 emissions, water management), social impacts (e.g., worker health and safety, impact on local communities), and governance (e.g., board diversity, corporate ethics, risk management).

Recent literature has begun to explore the link between ESG performance and firm value. Numerous studies suggest that strong ESG performance reduces the cost of capital, mitigates operational and reputational risks, and improves the long-term resilience of the firm. In the energy sector, which is highly exposed to regulatory and environmental risks, the integration of ESG criteria into decision-making and innovation processes has become an essential condition for survival and growth. The recent surge of mandatory disclosure regulations—such as the European Corporate Sustainability Reporting Directive (CSRD) and the EU Taxonomy—has further institutionalized ESG metrics, transforming them from voluntary signals into binding obligations that profoundly affect corporate strategy and reporting infrastructures.

3 Methodology

To rigorously and objectively map the intersection of energy innovation, circular economy, and ESG factors, this study adopts an approach based on quantitative bibliometrics. Bibliometrics allows for the analysis of large volumes of scientific literature, reducing the subjective biases typical of narrative reviews and providing a visual and statistical representation of the intellectual structure of a research field. The combination of performance analysis, science mapping, and network analysis offers a robust methodological foundation for capturing both the breadth and the depth of the literature.

3.1 Data Collection and Search Strategy

Data were extracted from the Web of Science (WoS) Core Collection database. The choice of WoS is motivated by its high quality, the rigorous selection of indexed journals, and the reliability of the provided metadata (such as cited bibliographies), which are essential for co-citation analyses.

The search strategy was constructed using a complex Boolean query, structured to intercept three conceptual macro-areas:

- Reference Sector: terms related to the energy context (e.g., "Energy sector", "Utilities", "Renewable Energy", "Power generation").
- Innovation and Transition: terms related to technological and process changes (e.g., "Smart Grid", "Circular Economy", "Digitalization", "Technological Innovation").
- Performance and Outcomes: terms associated with impact measurement and acceptance (e.g., "ESG", "Environmental Social Governance", "Firm Value", "Social Acceptance", "Sustainability Performance").

The search period was set from January 1, 2015, to January 31, 2026. The starting year, 2015, was strategically chosen as it coincides with two global milestones: the signing of the Paris Agreement on climate and the adoption of the 2030 Agenda for Sustainable Development by the United Nations, events that catalyzed academic and corporate attention on energy sustainability issues. The end of the period (January 2026) ensures the inclusion of the most recent and cutting-edge publications, capturing the latest evolutions in regulatory frameworks, technological breakthroughs, and managerial practices.

3.2 Selection and Filtering (PRISMA Protocol)

The selection process followed a logic similar to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) protocol. The initial extraction produced a total of 468 documents. Subsequently, rigorous inclusion and exclusion criteria were applied:

- Document Type: Only original research articles (articles) and literature reviews (reviews) were included, excluding editorials, book chapters, and conference proceedings to ensure the highest quality and the completion of the peer-review process.
- Language: The search was limited to documents written in English, considered the lingua franca of international scientific communication.

Following this refinement process, the final dataset under analysis stabilized at 395 documents. Each record was downloaded in plain-text BibTeX-compatible format, including full metadata such as authors, affiliations, abstracts, keywords (both authors' and KeyWords Plus), references, and citation counts.

3.3 Analysis Tools

The bibliometric analysis was conducted using Bibliometrix, an open-source package developed in the R programming environment, and its graphical interface Biblioshiny. This tool was selected for its versatility, transparency, and comprehensiveness. The techniques employed include:

- Performance analysis: to evaluate the productivity of authors, countries, and journals, alongside their citation impact.
- Keyword co-occurrence analysis: to map the conceptual structure and thematic evolution of the field.
- Co-citation analysis and historiographic mapping: to reconstruct the intellectual roots and evolutionary paths of the literature, identifying both seminal works and turning points.
- Collaboration network analysis: to visualize patterns of cooperation among authors, institutions, and countries.

4 Descriptive Results

The descriptive analysis of the dataset provides a fundamental overview of the health, size, and demographics of research in this domain.

4.1 Scientific Production and Growth Rate

The results reveal that the research field at the intersection of energy, innovation, and ESG is extremely dynamic. The annual growth rate of scientific production stands at 8.28%. This upward trend reflects the growing regulatory pressure (such as the European directives on sustainability reporting) and the interest of institutional investors in green finance. It also signals the maturation of a research community that increasingly sees the field as a legitimate and rewarding venue for academic contribution.

A particularly relevant finding concerns the impact of the research: the average number of citations per document is 26.31. This is an exceptionally high value, especially considering that the average age of the analyzed documents is only 3.7 years. This indicates that the articles published in this area are quickly read, recognized, and cited by the scientific community, confirming the "freshness" and urgency of the debate on sustainable energy transitions.

4.2 Source Analysis (Scientific Journals)

The profoundly interdisciplinary nature of this line of research is confirmed by the analysis of publication sources. The most relevant journals do not belong to a single sector (e.g., only engineering or only finance) but cover a broad spectrum of disciplines that combine environmental sciences, management, and technology.

The top three journals by number of publications are:

- Sustainability (43 articles): An open-access journal that serves as a main platform for cross-cutting studies on economic, environmental, and social sustainability.
- Energies (36 articles): Focused more on techno-economic aspects, energy technologies, and policies.
- Journal of Cleaner Production (22 articles): A very high-impact journal, historically linked to environmental engineering, circular economy, and sustainable production practices.

The concentration of articles in these three journals (about 25% of the total dataset) suggests the existence of a consolidated core of publications that drive the academic discourse on sustainable energy innovation. At the same time, the long tail of less concentrated venues confirms the field's multidisciplinary nature, where

contributions appear in journals as diverse as Renewable and Sustainable Energy Reviews, Business Strategy and the Environment, and Research Policy.

4.3 Geographical Distribution and Global Impact

The geographical analysis of scientific production shows a clear polarization, largely reflecting the global geopolitical and economic dynamics related to energy.

China absolutely dominates the research landscape, both in terms of production volume (262 articles) and in terms of impact (1,752 total citations). This primacy is not accidental but is a direct reflection of massive Chinese state policies aimed at dominating the supply chains of renewable energies (solar, wind, batteries) and managing the complex transition from coal to a low-carbon economy.

Italy ranks second, with 101 articles and 774 citations. This excellent result can be attributed to Italy's historical leadership in Europe in the field of smart grids (thanks to the early roll-out of smart meters) and a strong academic tradition in circular economy and waste management studies.

Following are India (84 articles) and the United Kingdom (82 articles). An anomalous and interesting finding emerges when analyzing Switzerland: although not in the top 4 for the number of publications, it ranks third globally for the number of citations (638). This indicates that the research produced in Switzerland, while quantitatively lower, possesses a disproportionately high quality and influence, likely due to the presence of excellent research institutes and international organizations based in Geneva and Zurich, which often act as bridges between academia, policy circles, and global governance bodies.

4.4 Author Profile and Collaboration Dynamics

The complexity of the energy "twin transition" requires skills that a single researcher rarely possesses. This is clearly reflected in the collaboration metrics. The dataset involves a total of 1,581 authors, with a very high collaboration index of 4.14 co-authors per document.

Furthermore, 38.23% of the documents feature forms of international co-authorship. Research on energy and ESG is therefore an intrinsically collective and global enterprise. Materials scientists, electrical engineers, environmental economists, and management scholars increasingly find themselves having to join forces to address systemic challenges, from the technical feasibility of a biogas plant to its acceptance by the local community and its impact on the ESG rating of the financing company. This trend mirrors the broader recognition that sustainability problems are by their nature "wicked," requiring interdisciplinary teams and methodological pluralism to be adequately addressed.

5 Conceptual Structure and Thematic Evolution

To understand what the scientific literature concretely discusses and how these topics have evolved over time, an analysis based on keywords (both Author Keywords and database-generated Keywords Plus) was conducted.

5.1 Keyword Co-occurrence Analysis

The co-occurrence network maps the keywords that jointly appear in the same articles, revealing the main thematic clusters. The analysis identified three fundamental macro-clusters:

- Strategic-Financial Cluster (ESG and Corporate Value): Dominated by terms such as "ESG performance", "Corporate Governance", "Financial performance", and "Firm value". This cluster highlights how the energy transition is now framed not only as an ethical or regulatory obligation but as a critical factor for mitigating financial risk and attracting capital.
- Technological-Operational Cluster (Innovation and Efficiency): Characterized by keywords such as "Renewable energy", "Smart grids", "Digitalization", "Artificial Intelligence", and "Energy efficiency". Here emerges the enabling role of digital technologies (the so-called Twin Transition, digital and green) in optimizing power grids and integrating intermittent renewable sources.
- Systemic-Environmental Cluster (Circular Economy and Decarbonization): Centered on "Circular economy", "Carbon neutrality", "Life cycle assessment (LCA)", and "Waste management". This group reflects the shift from a linear approach to circular business models within the utilities sector.

5.2 Thematic Map and Evolutionary Dynamics (2015-2026)

Using the Thematic Map (which positions themes on two axes: centrality/relevance and density/development), it is possible to trace the evolution of the debate:

- Motor Themes (High centrality and high density): The integration between "ESG" and "Financial Performance" remains the driving force of current research. Energy companies are the main testing ground to demonstrate that sustainability generates economic value.
- Basic Themes (High centrality, low density): "Renewable energy" and "Corporate Social Responsibility (CSR)" are positioned here. They are foundational but now mature concepts; CSR, in particular, is being progressively replaced by the ESG acronym, which offers more quantifiable metrics.
- Emerging Themes (Low centrality, low density): In the last three years (2023-2026), themes related to "Digital twin technology", "Green Hydrogen", and "Scope 3 emissions" have strongly emerged. This indicates that the research frontier is shifting towards measuring emissions along the entire value chain and using virtual simulations to optimize energy assets.

6 Intellectual Structure

The analysis of the intellectual structure allows us to identify the theoretical "roots" upon which today's research is founded, through the study of co-citations (two documents cited together by a third document).

6.1 Co-citation Analysis

The author co-citation network reveals that current literature relies on two distinct theoretical pillars that are merging:

On one hand, theorists of Institutional Economics and Stakeholder Theory (e.g., Freeman, Porter), used to explain the regulatory and social pressures driving energy companies to adopt ESG practices. On the other hand, scholars of Innovation Systems and Socio-Technical Transitions (e.g., Geels, Schot), fundamental for understanding how new technologies (e.g., smart grids, advanced storage) manage to overcome the barriers of the traditional fossil fuel market. The progressive convergence of these two streams reflects the recognition that technological feasibility and institutional acceptability are inseparable dimensions of any successful energy transition.

6.2 Historiographic Mapping

The historiographic analysis shows a clear turning point around 2019-2020. Before this two-year period, literature on energy innovation and sustainable finance traveled on parallel tracks. Starting in 2020, with the acceleration of the European Green Deal and the growing awareness of post-pandemic climate risks, a "merging" of citation paths is noted: technological articles begin to heavily cite financial studies and vice versa, sanctioning the birth of a truly interdisciplinary field of study. This convergence is also visible at the methodological level, with quantitative econometric techniques increasingly applied alongside qualitative case studies and engineering simulations.

7 Collaboration Networks

As anticipated in the descriptive results, the high complexity of the topic requires strong teamwork at a global level. The analysis of Collaboration Networks provides interesting details on how researchers interact.

7.1 Country Collaborations

The map of international collaborations is dominated by a strong China-United States and China-Europe (particularly the UK and Italy) axis. Despite global geopolitical tensions in the energy sector, academic cooperation remains solid. China, a leader in the production of hardware technologies (solar panels, batteries), collaborates intensely with European and North American institutions, which bring advanced expertise in circular business models, ESG regulatory frameworks, and climate finance. This pattern of complementarity helps explain why scientific cooperation has remained relatively resilient even during phases of trade tension and political friction.

7.2 Institutional Collaborations

At the institutional level, polytechnic universities (e.g., Politecnico di Milano, Tsinghua University, MIT) act as central "hubs" in the network. It is interesting to note a growing presence of co-authorships between universities and corporate research centers or international institutions (such as the World Bank or the International Energy Agency). This phenomenon of academic Open Innovation demonstrates that research in this field is strongly oriented towards practical application and supporting policy-making, blurring the traditional boundaries between scientific inquiry and operational decision-making.

8 Managerial and Policy Implications

The results emerging from this bibliometric analysis offer concrete insights for both corporate decision-makers and policy makers involved in the energy transition.

8.1 Managerial Implications

ESG as a strategic lever, not just compliance: The literature clearly demonstrates that ESG performance is closely correlated with corporate value and a reduction in the cost of capital. Managers of utilities and energy companies must integrate ESG criteria directly into technological innovation processes, shifting from a reactive approach (based on mere reporting) to a proactive one (based on sustainable design or eco-design). This shift requires a corresponding evolution in performance management systems, with executive incentives increasingly tied to ESG outcomes alongside financial targets.

Adoption of Circular Economy models: The integration of circularity into energy supply chains (for example, the recycling of lithium batteries or the reuse of materials from decommissioned wind turbine blades) emerges as a fundamental strategy to mitigate risks associated with the price volatility of critical raw materials. In a geopolitical environment characterized by supply chain disruptions and resource nationalism, circular strategies represent a form of strategic insurance.

Leveraging the "Twin Transition": The thematic analysis highlights the crucial importance of digitalization. Managers should invest in technologies such as Artificial Intelligence, IoT, and Digital Twins not only to optimize the operational efficiency of grids but also to transparently track and report environmental and social data in real-time. The credibility of ESG disclosures will increasingly hinge on the robustness of the underlying data infrastructure.

8.2 Policy Implications

Harmonization of metrics: The proliferation of different ESG standards creates confusion and favors greenwashing. Governments and regulatory bodies must accelerate towards a global standardization of sustainability metrics, making data comparable and verifiable. International cooperation among standard-setters such as the ISSB, EFRAG, and the SEC is essential to avoid regulatory fragmentation.

Incentivizing systemic innovation: Public policies should not be limited to subsidizing a single technology (e.g., solar panels) but should incentivize the entire circular ecosystem, promoting industrial symbiosis and collaboration across different sectors. Cross-sectoral incentives, public-private partnerships, and innovation procurement can play a significant role in stimulating systemic transformation.

9 Conclusions

This study explored the evolution and intellectual structure of scientific research at the intersection of energy innovation, circular economy, and ESG factors, analyzing 395 documents published between 2015 and 2026.

The results confirm that we are facing a highly dynamic field of study (with an annual growth rate exceeding 8%) with a high impact. The analysis revealed a progressive shift in academic focus: while in the early years the debate was dominated by theoretical concepts related to Corporate Social Responsibility (CSR), today research is strongly oriented towards quantitative metrics (ESG), enabling technologies (digitalization), and transformative business models (circular economy).

Geographically, research is led by China (by volume) and Europe (with Italy and the UK at the forefront of systemic innovation), supported by solid international collaboration networks that withstand geopolitical fragmentation. From an intellectual perspective, we are witnessing the merging of two historically separate streams: the engineering of socio-technical transitions and sustainable corporate finance. This convergence not only enriches the conceptual depth of the field but also enhances its practical relevance, providing decision-makers with integrated frameworks rather than isolated technical or financial perspectives.

10 Limitations of the Study and Future Research Agenda

Like any research, this study has limitations. First, the analysis is based exclusively on the Web of Science database and English-language documents; the inclusion of other databases (such as Scopus or Dimensions) or "grey" literature (institutional reports, white papers, and industry analyses) could further enrich the picture. Second, bibliometric analysis provides a quantitative map, which always needs to be integrated with qualitative reviews to capture the methodological nuances of individual articles. Third, the keyword-based approach inevitably depends on authors' tagging choices, which may be inconsistent across journals and disciplines.

Based on the gaps identified in the thematic map, we propose the following agenda for future research:

The role of the "S" (Social) in ESG: Currently, the literature is heavily skewed towards the environmental and governance dimensions. Future studies are needed to investigate the social impact of the energy transition, such as the creation of "green jobs," the social acceptability of new infrastructures, and the fight against energy poverty (the so-called Just Transition).

Measurement of Scope 3 Emissions: With the evolution of regulations, the real challenge for energy companies will be measuring and reducing indirect emissions across the entire value chain (Scope 3). Research will need to develop models and frameworks to facilitate this complex reporting.

Impact of Artificial Intelligence on ESG: Although AI is an efficiency enabler, training models requires enormous amounts of energy. It will be crucial to investigate the paradox between the benefits of AI for smart grid management and its own environmental impact (the digital carbon footprint).

Detecting Greenwashing: With the growing financial pressure linked to ESG ratings, there is an urgent need for empirical studies that develop advanced methodologies to identify and measure greenwashing and circularwashing practices in the energy sector. The development of natural-language-processing tools, sentiment analysis algorithms, and forensic accounting techniques will be central in equipping investors, regulators, and civil society with reliable instruments to discriminate between substantive and rhetorical sustainability commitments.

11 Ethics declaration

This study is based on publicly available bibliographic information retrieved from indexed scientific databases and does not involve, at any stage, the participation of human subjects, animals, or personally identifiable information. As the research relies exclusively on secondary data drawn from peer-reviewed publications and metadata, prior approval from an institutional ethics review board was not required, in line with prevailing guidelines for desk-based bibliometric studies. The authors adhered to the general principles of research integrity, including reliability of the analysis, honesty in the reporting of methods and results, accurate citation of prior scholarship, and accountability for the entire research process. All sources have been duly acknowledged, no data have been fabricated or selectively omitted, and no conflict of interest exists that could have influenced the design or interpretation of the study. Should future extensions of this research involve the collection of primary data, the authors commit to obtaining the necessary ethical clearances from the competent committees, ensuring informed consent, voluntary participation, and full compliance with applicable data-protection regulations.

12 AI declaration

In the preparation of this manuscript, generative Artificial Intelligence tools were used only as auxiliary instruments for minor tasks such as linguistic refinement, grammatical checking, and the formal organization of references. Every output produced with AI assistance was critically reviewed and validated by the authors before inclusion in the final text.

References

Barney, J. (1991) “Firm Resources and Sustained Competitive Advantage”, *Journal of Management*, Vol. 17, No. 1, pp 99–120.

Bergek, A., Jacobsson, S., Carlsson, B., Lindmark, S. and Rickne, A. (2008) “Analyzing the functional dynamics of technological innovation systems: A scheme of analysis”, *Research Policy*, Vol. 37, No. 3, pp 407–429.

Cucchiella, F., D'Adamo, I. and Gastaldi, M. (2018) “Profitability analysis of small-scale plants for biomethane injection into the gas grid”, *Journal of Cleaner Production*, Vol. 184, pp 179–187.

Ellabban, O., Abu-Rub, H. and Blaabjerg, F. (2014) “Renewable energy resources: Current status, future prospects and their enabling technology”, *Renewable and Sustainable Energy Reviews*, Vol. 39, pp 748–764.

Geels, F.W. (2002) “Technological transitions as evolutionary reconfiguration processes: a multi-level perspective and a case-study”, *Research Policy*, Vol. 31, No. 8-9, pp 1257–1274.

Geissdoerfer, M., Savaget, P., Bocken, N.M.P. and Hultink, E.J. (2017) "The Circular Economy – A new sustainability paradigm?", *Journal of Cleaner Production*, Vol. 143, pp 757–768.

Ghosh, S. (2022) "Strategic sourcing model for green supply chain management: An insight into automobile manufacturing units in India", *Benchmarking: An International Journal*, Vol. 29, No. 10.

Markard, J., Wirth, S. and Truffer, B. (2016) "Institutional dynamics and technology legitimacy – A framework and a case study on biogas technology", *Research Policy*, Vol. 45, No. 1, pp 330–344.

Contextual–RFM (C-RFM): a Multimodal Framework for the Fusion of Transactional and Semantic Data Through LLMs for Interactive Customer Segmentation

Davide Madonna, Simone Malacaria, Andrea De Mauro
Luiss University, Rome, Italy

Abstract

This study examines the evolution of customer segmentation from rule-based systems to multimodal AI architectures, emphasizing the necessity for transparent and context-aware CRM strategies. Despite the widespread adoption of RFM models, organizations often struggle to synthesize transactional metrics with multifaceted qualitative feedback, creating a numerical paradox that this research addresses by capturing the authentic voice of the customer within high-dimensional analytical spaces. This paper presents the Contextual-RFM (C-RFM) framework, an innovative multimodal architecture aimed at bridging the representational divergence between quantitative metrics and qualitative feedback.

The research operationalizes a comprehensive five-phase pipeline to ensure methodological transparency and replicability: (1) establishment of a statistical baseline, (2) narrative synthesis for tabular serialization, (3) multimodal vectorization utilizing advanced embeddings, (4) semantic analysis through Latent Dirichlet Allocation (LDA) and (5) implementation of the C-RFM Expert chatbot. This entire workflow is implemented using KNIME Analytics and Langflow, providing a robust and replicable process for both scholars and practitioners. A key distinction of the framework is its early-fusion approach. Unlike traditional late-fusion models that integrate data streams only at the final analysis stage, early-fusion retains critical cross-modal correlations during the initial processing phase, effectively addressing the semantic opacity common in traditional numerical-only datasets.

Empirical findings indicate that the C-RFM framework adeptly organizes linguistic complexity, evidenced by a significant improvement in structural validity when transitioning from purely transactional environments to topic-validated contexts. Moreover, this study introduces the Discrepancy Detection Rate (DDR) as a novel performance metric to quantify the misalignment between transactional behaviour and qualitative feedback, thereby identifying critical churn risks that conventional quantitative analyses often overlook.

The primary contribution of this research lies in the synergy achieved between human expertise and Artificial Intelligence. By deploying dynamic persona chatbots grounded in Retrieval-Augmented Generation (RAG), the framework facilitates a shift from opaque analytical outputs to actionable, customized recovery strategies. This glass-box methodology transforms intricate numerical complexities into auditable business narratives, empowering decision-makers to transcend statistical constructs and foster a deeper, dialogue-based understanding of consumers. The C-RFM framework represents a significant advancement in human-centred AI, promoting strategic marketing intelligence that is not only transparent but also finely tuned to the authentic voices of real-world consumer archetypes, thereby paving the way for the next generation of customer-centric strategies.

Keywords: C-RFM Framework, Large Language Models, Explainable AI, Multimodal Data Fusion, Human-Centred CRM

Paper type: Academic Research Paper

1 Introduction

Artificial Intelligence is fundamentally transforming the Customer Relationship Management (CRM) landscape by enabling a significant shift from standardized marketing approaches to highly personalized interactions (Naderipour and Hosseini Nasabian, 2025; Bukhari et al., 2022; Kandi and Basani, 2024). In this evolving framework, segmentation continues to play a crucial role, allowing organizations to effectively manage market density by identifying and isolating specific behavioural groups (Syahra, Fadlil and Yuliansyah, 2025; Akande, Asani and Dautare, 2024). Historically, the theoretical foundation of this field has been the Recency, Frequency, and Monetary (RFM) model. However, its evolution from rule-based systems to unsupervised machine learning, notably K-means clustering, has exposed a persistent numerical paradox: a reliance on purely transactional data that lacks the contextual depth necessary to capture the multifaceted nuances of the modern consumer journey (Lewaa, 2023; Gopakumar et al., 2021).

The technical roots of this limitation lie in traditional One-Hot Encoding (OHE), which often suffers from the dimensionality curse and statistical instability (Chung, 2024). More critically, OHE induces a condition of semantic opacity within datasets. As Bayrak (2022) and Liang (2025) highlight, by treating all categories as mathematically independent, this method fails to recognize latent behavioural similarities between consumer

actions. Contemporary analytical frameworks are transitioning toward deep semantic embeddings to resolve these structural flaws. Smutek et al. (2024) utilize encoder-decoder structures to compress high-dimensional inputs into dense vectors where numerical proximity reflects semantic relationships. Validating these learned representations presents a significant challenge: while the Silhouette score is traditionally utilized as the primary metric for cluster validation, assessing structural quality by measuring both intra-cluster cohesion and inter-cluster separation (Gagolewski, Bartoszuk and Cena, 2021; Shutaywi and Kachouie, 2021), its effectiveness is often diminished in embedding spaces. The dimensionality curse and latent space sparsity make data points appear nearly equidistant, rendering traditional distance-based validation unreliable (Simpson et al., 2025). Researchers have increasingly adopted Latent Dirichlet Allocation (LDA) as a robust alternative to overcome these geometric limitations, thereby overcoming geometric limitations while demonstrating remarkable stability even when semantic relationships are intricate or data is sparse (Sidek and Ahmad, 2026).

Treating transactional data as a linguistic structure (Woltmann, Thiele and Lehner, 2018) positions Large Language Models (LLMs) as the next transformative frontier, leveraging their emergent abilities for data synthesis and optimization-based feature engineering (Nam et al., 2024). Despite this potential, LLM integration into analytical workflows faces a significant numerical paradox, where standard Byte-Pair Encoding (BPE) tokenizers fragment digits into subunits without inherent meaning, which, as Penttala (2025) states, strips numerical values of their quantitative significance. Hegselmann et al. (2023) demonstrate that serializing tabular data into natural language strings allows models to leverage their extensive prior knowledge to mitigate these errors, effectively interpreting transactional records as coherent behavioural narratives. This trajectory culminates in multimodal fusion strategies, which merge structured metrics with unstructured insights through sophisticated transformer-based architectures to capture cross-modal interactions and improve predictive transparency (Hangloo and Arora, 2025). Yet, even as these models enhance analytical depth, the transition from opaque outputs to actionable managerial trust remains a critical hurdle. This difficulty is primarily driven by the inherent lack of transparency in high-dimensional clustering, which frequently generates skepticism among strategic decision-makers (Wang and Ding, 2024). In critical CRM contexts, Explainable AI (XAI) offers a vital methodological framework for converting opaque computations into narratives that are comprehensible to humans (Patil, 2024). Beyond merely achieving predictive accuracy, XAI employs post-hoc tools such as SHAP and counterfactual analysis to uncover behavioural drivers, thereby bridging the gap between statistical constructs and strategic marketing intelligence while ensuring ethical and accountable insights (Almtrf, 2025). This integration is exemplified in dynamic persona chatbots, which transform the interaction between decision-makers and consumer archetypes from static descriptions to engaging, real-time dialogues (Salminen et al., 2026). By leveraging text embeddings models and direct LLM prompting to create persona chatbots, these agents accurately simulate consumer preferences, fostering a trust-based synergy between humans and AI to address complex marketing challenges (Li, Liu and Yu, 2025; Huang, Markovitch and Stough, 2024). This paper introduces the Contextual-RFM (C-RFM) framework to reconcile quantitative rigor with semantic interpretation. In addressing the heterogeneity gap that arises when comparing distinct representation spaces (Hangloo and Arora, 2025), C-RFM employs an early-fusion approach. By integrating transactional and semantic data during the initial preprocessing phase, the framework produces robust embeddings that overcome both the semantic limitations of traditional models and the numerical paradox associated with tokenization (Avnat et al., 2024). This methodology equips LLMs with a deeper contextual understanding while alleviating the information loss and managerial distrust often linked to high-dimensional clustering. By combining XAI with interactive chatbots, C-RFM enables an interactive data interpretation, addressing the following research questions:

- *RQ1: Which methodological approaches prove most effective for validating the multimodal segmentation of integrated transactional and semantic data?*
- *RQ2: How can an LLM-embedding framework, implemented through Explainable AI and interactive persona chatbots, effectively bridge the divide between opaque analytical outputs and actionable, transparent strategies?*

2 Methodology

The C-RFM framework is operationalized through a comprehensive five-phase pipeline, which serves as a procedural structure for converting the theoretical early-fusion strategy into a replicable analytical process.

2.1 Phase 1: Statistical baseline

This phase establishes a quantitative benchmark utilizing the Olist Brazilian E-Commerce dataset. A refined sample of 4,783 profiles was processed in KNIME Analytics, incorporating relational data from customers, orders, payments, and reviews. Procedural filters were applied to remove low-density reviews, while stratified sampling

was employed to maintain variance within high-value segments. A baseline RFM analysis was performed using K-means clustering, optimized through a two-stage methodology, which identified three segments as the optimal configuration. Cluster validation was carried out using the Silhouette score, resulting in a mean coefficient of 0.474. Although this outcome is considered acceptable, it highlights the structural ambiguities and boundary overlaps present in raw transactional data.

2.2 Phase 2: Narrative synthesis

The limitations of traditional tokenization are mitigated by replacing statistical grouping with row-level textual representations. The dimensions of RFM were redefined through quantile binning into semantic categories, transforming numerical variance into contextually significant data for LLMs. As suggested by Hegselmann et al. (2023), this serialization enables models to utilize extensive prior knowledge to interpret transactional records as coherent behavioural narratives. By integrating these categories with the verified customer perspective, the framework preserves cross-modal correlations prior to high-dimensional processing.

2.3 Phase 3: Multimodal vectorization

This phase implements the early-fusion strategy by embedding synthesized narratives into a high-dimensional space using the OpenAI *text-embedding-3-small* model. To avoid data leakage and facilitate the identification of new latent relationships, narratives were cleansed of Phase 1 RFM labels. Dimensionality was managed through Principal Component Analysis (PCA), which retained 100 components, accounting for 76% of the variance, followed by K-means clustering with three clusters to ensure stable segmentation. The analysis indicated a notable transition between transactional and semantic profiles. As illustrated in the migration matrix (Figure 1), RFM clusters are predominantly aligned with embedding cluster 0, an occurrence attributed to the specific feedback distributions found within the Olist dataset. Despite this skewness, cross-tabulation analysis revealed a p-value of 0.001, indicating that high-value transactional segments are distributed across diverse semantic contexts. This finding highlights the limitations of relying solely on raw numerical data and emphasizes the need for enhanced semantic depth (Li, Liu and Yu, 2025).

Frequency Row Percent	0	1	2	Total
At-Risk Customers	998	314	533	1.845
	54,0921%	17,019%	28,8889%	
Lost Customers	1.096	386	531	2.013
	54,4461%	19,1754%	26,3785%	
Top Customers	449	166	310	925
	48,5405%	17,9459%	33,5135%	
Total	2.543	866	1.374	4.783

Figure 1 – Migration matrix of RFM and embeddings clusters.

2.4 Phase 4: Semantic analysis

Topic modelling using LDA revealed behavioural communities within the high-dimensional vector space. Following a text mining workflow designed to minimize linguistic noise and an evaluation of the perplexity index, three distinct themes emerged. Labels for these topics were assigned based on the semantic relevance of the most representative terms identified during the modelling process (Figure 2). Topic 0 was labelled as *delivery excellence and reliability*, indicating positive logistical experiences; topic 1 was characterized as *product discrepancy and material quality*, which highlighted concerns regarding the physical attributes of products; and topic 2 was designated as *logistic and administrative friction*, focusing on post-sale and bureaucratic challenges.

Topic id String	Concatenate(Term) String
topic_0	recommend, quality, ahead, packaged, price, earlier, expectations, quickly, advertised, correct
topic_1	quality, missing, box, photo, material, original, watch, bad, little, fit
topic_2	delay, invoice, wrong, response, email, requested, day, refund, address, exchange

Figure 2 – Representative terms extracted for each topic

The assignment of narrative identities to the embedding clusters was informed by their distribution across these themes (Figure 3). As a result, cluster 0 was classified as *effortless brand ambassadors*, cluster 1 as *quality-focused skeptics*, and cluster 2 as *administrative hostages*. The structural integrity of these segments was further substantiated through unigram and bigram validation, culminating in the development of a high-fidelity customer narrative.

Semantic Cluster String	Delivery Excellence & Reliability... Number (Integer)	Logistic & Administrative Friction... Number (Integer)	Product Discrepancy & Material Quality Number (Integer)
Administrative Hostages	411	805	158
Effortless Brand Ambassadors	2096	234	213
Quality-Focused Skeptics	165	130	571

Figure 3 – Distribution of embeddings clusters across semantic topics

This format serves as the foundational contextual input for the persona chatbot, integrating transactional data, experience scores, and semantic analysis into a cohesive behavioural profile (Figure 4).

```
"CUSTOMER PROFILE: This user is a '", $RFM Cluster$, "' (categorized as: ", $Recency Label$, ", ", $Frequency Label$, ", and ", $Monetary Value Label$, "). ",
"TRANSACTIONAL DATA: Last purchase: ", string($Recency$), " days ago. Total activity: ", string($Frequency$),
" orders with a total lifetime value of ", string($Monetary Value$), " Reais. ",
"EXPERIENCE DATA: Review Score: ", string($Review Score$), "/5 stars. Feedback: '", $Review Text$, "'. ",
"SEMANTIC ANALYSIS: Cluster: '", $Semantic Cluster$, "'. Topic: '", $Topic Name$, "'. ",
"CONTEXTUAL KEYWORDS: ", $Topic Keywords$, ". "
```

Figure 4 – Prompt engineering template for customer stories

2.5 Phase 5: Implementation of the C-RFM Expert

The final phase focuses upon the deployment of the C-RFM Expert through a RAG architecture, which serves as the primary interface for narrative explainable AI. The implementation is structured around three core processes: knowledge base construction, orchestration via Langflow, and agentic deployment.

The construction of the knowledge base began with 1,000 representative vectorized narratives. A distributed processing approach utilizing a LLMs was employed to conduct a discrepancy analysis, aligning transactional metrics with sentiments derived from qualitative feedback. This methodology facilitates the categorization of customers into semantic clusters and the development of recovery plans that are specifically tailored to the unique behavioural nuances identified within each profile narrative.

System orchestration was executed in Langflow, utilizing Chroma DB for long-term multidimensional storage and a *gpt-4o-mini* agent to independently search the vector store. The resulting C-RFM Expert persona operates under a system prompt designed to identify friction points and rationalize strategic interventions. This approach guarantees that strategic recommendations are both auditable and grounded in empirical customer narratives, effectively converting numerical complexity into actionable business intelligence.

3 Results and Discussion

The following sections outline the empirical findings of the C-RFM framework, emphasizing the validation of the multimodal approach and its strategic advantages compared to conventional transactional models. By integrating quantitative metrics with generative insights, the results illustrate how the numerical paradox can be effectively addressed through both structural and semantic alignment.

3.1 Quantitative performance and validation

The validation of the C-RFM framework employed a multi-faceted statistical approach, incorporating the Silhouette score to assess segment cohesion, the Adjusted Rand Index (ARI) to measure model novelty, and the Discrepancy Detection Rate (DDR), a novel metric specifically developed and implemented through this study, to quantify the business value of uncovering latent attrition risks.

3.1.1 Multimodal structural validation

A significant finding emerged when traditional RFM clusters were projected onto the PCA-derived semantic space, resulting in a Silhouette score of 0.000. This outcome suggests that transactional segments are statistically indistinguishable from random noise when assessed through customer feedback, thereby providing empirical evidence that purely quantitative models lack semantic richness. Within this conventional framework, RFM categories do not share a unified linguistic or experiential narrative, highlighting that high spending does not necessarily equate to customer satisfaction and that transactional loyalty can often obscure underlying administrative challenges. In contrast, the C-RFM framework effectively organized this linguistic complexity. As illustrated Table 1, the transition from raw PCA space to probabilistic topic distributions significantly improved structural validity.

Table 1 – Comparative analysis of Silhouette scores across multimodal analytical spaces

Model/Space	Silhouette score	Interpretation
RFM clusters on PCA space	0.000	Total semantic opacity
C-RFM clusters on PCA space	0.161	Emergence of latent structure
C-RFM on topic matrix (LDA)	0.249	Probabilistic consolidation

3.1.2 Innovation and risk: ARI and DDR

The scientific novelty of the model was evaluated by calculating the Adjusted Rand Index (ARI) between the RFM clusters from Phase 1 and the semantic clusters from Phase 4. The primary aim of the ARI analysis was to assess the degree of similarity, or lack thereof, and potential redundancy between the established transactional model and the proposed multimodal framework (Sundqvist, Chiquet and Rigaill, 2023).

The computed ARI of 0.005 represents a significant finding, indicating that the two partitioning methods are stochastically independent. This near-zero value underscores a phenomenon of semantic opacity, suggesting an almost non-existent correlation between a customer's spending behaviour and their qualitative experiences. While an ARI close to 1 would indicate redundancy, this outcome confirms that the C-RFM framework reveals distinct behavioural dimensions, demonstrating its capacity to provide a non-redundant and unique strategic perspective.

Additionally, the framework presents the DDR as an innovative performance metric, which is established through the statistical cross-validation of transactional RFM segments and semantic topic clusters. This metric is defined by the following ratio:

$$DDR = \frac{|\{c \in Starget : \omega(c) \in Crisk\}|}{|Starget|}$$

- Starget: this refers to the group of customers identified within the highest transactional value segment.
- Crisk: this denotes the subset of semantic clusters that highlight critical friction points.
- ω(c): this is the semantic mapping function that assigns a customer, denoted as c, to a specific cluster based on their textual feedback.
- |·|: this operator indicates cardinality, representing the total number of individuals within the specified set.

Derived from the empirical values within the dataset, the computed DDR of 0.515 exposes a significant latent churn crisis, indicating that 51.5% of the company's most valuable customers are classified within high-risk clusters. This underlying threat underscores the fundamental limitations of conventional RFM analysis, which often fails to identify these critical churn risks (Lewaa, 2023). LDA topic modelling validates the internal differentiation within traditional RFM clusters, proving that customers are not a homogeneous entity but are instead fragmented across diverse experiential themes.

3.2 Managerial interpretability

The evolution from abstract clusters to persona chatbots signifies the peak of Human-Computer Interaction (HCI) (Salminen et al., 2026). By transforming statistical groupings into strategic archetypes, the framework effectively connects data science with actionable marketing strategies. The deployment of the C-RFM Expert agent illustrates that LLMs, when supported by a RAG vectorized knowledge base, can conduct discrepancy analysis and generate prescriptive strategies with high accuracy.

This interpretability is implemented within the Langflow playground, where the agent enables interactive segmentation. Unlike conventional dashboards that present static data lists, the agent empowers managers to query the database using context-sensitive criteria. A pertinent example of this functionality is illustrated by the following strategic inquiry:

"Identify a high-value customer in the database who had a negative experience with the last purchase and provide a brief report specifying their exact numerical Recency, Frequency, and Monetary values, their Semantic Cluster name, a quote of their actual Review Text, and a concise 3-step recovery plan to resolve the discrepancy between their high transactional value and negative feedback".

As demonstrated in Figure 5, the agent successfully retrieved a high-value profile, accompanied by feedback expressing significant frustration with delivery failures.

High-Value Customer Report

Customer Profile:

- **Recency:** 192 days since last purchase
- **Frequency:** 18 orders
- **Monetary Value:** 181.86 Reais
- **Semantic Cluster Name:** Administrative Hostages

Review Feedback:

- **Rating:** 1.0/5 stars
- **Quote:** "The delivery was severely lacking, leading to frustration and dissatisfaction."

Recovery Plan

1. **Immediate Resolution:**
   - Contact the customer to provide an update on their delivery status. If the product is confirmed lost or significantly delayed, offer a prompt refund or replacement. Ensure clear communication about the steps being taken to resolve the issue.
2. **Compensation Offer:**
   - As a gesture of goodwill, offer free expedited shipping on their next order and a 20% discount on their next purchase. This acknowledges their loyalty and addresses the logistical issues they faced.
3. **Follow-Up Communication:**
   - After resolving the issue, follow up with the customer to ensure their satisfaction with the resolution and to reinforce the commitment to improving their experience. This can help rebuild trust and encourage future purchases.

Figure 5 – C-RFM Expert prescriptive response

The output generated by the C-RFM agent represents a significant shift in analytical detail and segmentation effectiveness. While traditional RFM clustering provides a purely quantitative perspective on consumer value, remaining semantically opaque regarding the underlying causes of attrition, the C-RFM framework leverages the agent's ability to integrate transactional significance with detailed linguistic feedback.

By following the agent's reasoning, managers could obtain an auditable and evidence-based understanding of the proposed strategy. This glass-box methodology ensures that each recommendation is underpinned by a clear rationale, enabling decision-makers to transition from descriptive insights to prescriptive interventions in real-time (Patil, 2024). Such transparency effectively alleviates the managerial skepticism commonly associated with black-box automated systems (Wang & Ding, 2024).

3.3 Academic and practical implications

The results of this study provide a dual-layered contribution, connecting advanced data science with strategic marketing management.

3.3.1 Academic implications

This research resolves the numerical paradox by demonstrating that early-fusion textualization enables LLMs to preserve and interpret cross-modal correlations. The shift from a Silhouette score of 0.000 to 0.249 illustrates that semantic signals can be mathematically structured to enhance transactional data.

However, it is important to acknowledge the absolute value of these structural metrics. Although this transition signifies a measurable improvement, scores remain lower than traditional RFM baselines, aligning with literature on the dimensionality curse in high-dimensional embedding spaces. This structural compression arises because the Silhouette score on Euclidean distance, which often fails to capture true proximity within high-dimensional contexts where latent space sparsity complicates rigid partitioning (Simpson et al., 2025). The framework addresses this challenge by shifting from geometric rigidity to the probabilistic logic of LDA, supplemented by the ARI and DDR metrics. While ARI provides empirical proof of the framework's novelty by confirming the stochastic independence of transactional and semantic spaces, the DDR offers a non-geometric validation of the model's strategic utility. By modelling the relationship between documents and topics as a distribution, the framework navigates the inherent ambiguity and polysemy of natural language more effectively than traditional distance metrics. Thus, the observed spatial overlap should not be interpreted as a failure of the model, but rather as a reflection of the fuzzy and continuous nature of semantic meanings (Petukhova et al., 2024).

3.3.2 Practical implications

The C-RFM framework effectively addresses hidden customer attrition by shifting from aggregate, retrospective segmentation to proactive, hyper-personalized recovery strategies. While traditional RFM models often lack insight into the emotional and logistical challenges that accompany high-value transactions, this framework functions as a scalable tool that converts passive data points into dynamic, conversational intelligence.

Organizations can implement interventions that are both semantically and economically tailored by automating the identification of strategic archetypes. Such a transformation ensures that recovery efforts directly address the root causes of dissatisfaction, tackling specific friction points that numerical metrics alone may overlook. Additionally, the introduction of an agentic interface effectively mitigates the managerial skepticism. The framework establishes AI as an advanced strategic partner by integrating complex multidimensional insights into actionable and verifiable business narratives, thereby ensuring a human-centred and evidence-based response.

4 Conclusions and Future Research

The C-RFM framework operationalized in this study effectively reconciles the quantitative inconsistencies inherent in traditional segmentation methods with a transparent model of XAI through the implementation of an early-fusion architecture. Its primary contribution lies in the practical validation of this early-fusion approach, which minimizes information loss and generates integrated semantic embeddings, thereby preserving semantic richness.

In addressing RQ1, the study illustrates that a comprehensive validation strategy is crucial for effective integrated multimodal segmentation. By merging structured narrative synthesis with LDA topic modelling and structural metrics, the framework identifies behavioural segments that are both statistically robust and semantically distinct. This methodology ensures that the resulting clusters transcend mere mathematical constructs, representing meaningful and interpretable consumer archetypes.

Concerning RQ2, the framework bridges the divide between opaque analytical outputs and actionable strategies through an agentic RAG architecture. The C-RFM Expert persona exemplifies how interactive chatbots, rooted in validated narratives, can automate discrepancy assessments and generate customized recovery plans. This shift from static categorizations to dynamic dialogue guarantees that marketing strategies are contextually transparent, data-driven, and auditable by human stakeholders.

Additionally, the introduction of the DDR signifies a foundational innovation for identifying latent customer attrition. As a novel metric, it serves as a crucial starting point for iterative development, facilitating further refinement as it is applied across various organizational contexts to quantify the misalignment between transactional value and consumer sentiment.

Future research could enhance the C-RFM framework by tackling current challenges related to data acquisition and the computational demands associated with commercial LLM interface. By incorporating socio-demographic variables, subsequent studies can further enrich the depth and detail of customer narratives, yielding more comprehensive contextual insights. This framework can also serve as a significant advancement

in human-centred CRM, fostering a more ethical and accountable approach that aligns data-driven accuracy with the complex realities of human behaviour.

5 Ethics Declaration

No ethical approval was required for this study.

6 AI Declaration

LLMs including Notebook LM were employed for content extraction, while Gemini was employed for linguistic refinement. Gemini specifically assisted in proofreading, structural clarification, and the formalization of the DDR formula, which is firmly grounded in the author's conceptual framework and empirical data. All substantive content, interpretations, and analytical contributions remain the author's original work. These tools provided purely editorial and formalization support, without influencing the intellectual or methodological foundations of the research.

References

Akande, O. N., Asani, E. O. and Dautare, B. T. (2024) "Customer segmentation through RFM analysis and K-means clustering: Leveraging data-driven insights for effective marketing strategy", *Ceddi Journal of Information System and Technology (JST)*, Vol. 3, No. 1, pp 14-25. https://doi.org/10.56134/jst.v3i1.81

Almtrf, A. (2025) "Integrating Explainable Ai (Xai) Into Decision Support Systems: A Framework for Enhancing Transparency and Trust in Managerial Decision-Making", *International Journal of Managerial Studies and Research (IJMSR)*, Vol. 13, No. 9, pp 9-22. https://doi.org/10.20431/2349-0349.1309002

Avnat, E., Levy, M., Herstain, D., Yanko, E., Ben Joya, D., Tzuchman Katz, M., Eshel, D., Laros, S., Dagan, Y., Barami, S., Mermelstein, J., Ovadia, S., Shomron, N., Shalev, V. and Abdulnour, R.-E.E. (2025). "Performance of large language models in numerical vs. semantic medical knowledge: Benchmarking on evidence-based Q&As". *Artificial Intelligence in Medicine*, 151, 102943. https://doi.org/10.2196/64452

Bayrak, A. T. (2022) "An application of Customer Embedding for Clustering", *2022 IEEE International Conference on Data Mining Workshops (ICDMW)*, pp 914–920. https://doi.org/10.1109/ICDMW58026.2022.00019

Bukhari, T. T., Oladimeji, O., Etim, E. D. and Ajayi, J. O. (2022) "Customer Lifetime Value Prediction Using Gradient Boosting Machines", *Gyanshauryam*, *International Scientific Refereed Research Journal*, Vol. 5, No. 4, pp 488–506. https://gisrrj.com/paper/GISRRJ225427.pdf

Chung, S. (2024) "Limitations and Problems of Binarization with One Hot Encoding", *CIS 660 Data Mining with Advanced Machine Learning*. Cleveland State University.

Gagolewski, M., Bartoszuk, M. and Cena, A. (2021). "Are cluster validity measures in fact valid?" *Information Sciences*, 581, 620–636. https://doi.org/10.1016/j.ins.2021.10.004

Gopakumar, A. B., Krishnan, G., Rozario, N. L., Nazreth, R. B. and Anilkumar, P. (2021) "Customer segmentation using K - Means Algorithm", *International Journal of Innovative Research in Technology (IJIRT)*, Vol. 8, No. 7, pp 127–132. https://ijirt.org/publishedpaper/IJIRT154245_PAPER.pdf

Hangloo, S. and Arora, B. (2025) "Multimodal fusion techniques: Review, data representation, information fusion, and application areas", *Neurocomputing*, Vol. 649, Art. 130827. https://doi.org/10.1016/j.neucom.2025.130827

Hegselmann, S., Buendia, A., Lang, H., Agrawal, M., Jiang, X. and Sontag, D. (2023) "TabLLM: Few-shot classification of tabular data with large language models", in Ruiz, F., Dy, J. and van de Meent, J. W. (eds.) *Proceedings of the 26th International Conference on Artificial Intelligence and Statistics (AISTATS 2023)*, Proceedings of Machine Learning Research, Vol. 206, pp. 5549–5581. https://proceedings.mlr.press/v206/hegselmann23a.html

Huang, D., Markovitch, D. G. and Stough, R. A. (2024) "Can chatbot customer service match human service agents on customer satisfaction? An investigation in the role of trust", *Journal of Retailing and Consumer Services*, Vol. 76, Art. 103600. https://doi.org/10.1016/j.jretconser.2023.103600

Kandi, A. and Basani, M. A. R. (2024) "Personalization and Customer Relationship Management in AI-Powered Business Intelligence", *International Journal for Research in Applied Science & Engineering Technology (IJRASET)*, Vol. 12, No. 11, pp. 714-715. https://doi.org/10.22214/ijraset.2024.65159

Lewaa, I. (2023) "Customer Segmentation Using Machine Learning Model: An Application of RFM Analysis", *Journal of Data Science and Intelligent Systems*. https://doi.org/10.47852/bonviewJDSIS32021293

Li, Y., Liu, Y. and Yu, M. (2025) "Consumer segmentation with large language models", *Journal of Retailing and Consumer Services*, Vol. 82, Art. 104078. https://doi.org/10.1016/j.jretconser.2024.104078

Liang, Z. (2025). "Efficient representations for high-cardinality categorical variables in machine learning". In *Proceedings of the 2025 IEEE International Conference on Advanced Machine Learning and Data Science (AMLDS 2025).* IEEE. https://doi.org/10.1109/AMLDS63918.2025.11159409

Naderipour, M. and Hosseini Nasabian, H. (2025). "A comprehensive RFM-A framework: Integrating age for enhanced customer segmentation and marketing strategies". *Transactions on Soft Computing*, 1(1), 14–26. https://doi.org/10.48314/tsc.v1i1.34

Nam, J., Kim, K., Oh, S., Tack, J., Kim, J. and Shin, J. (2024) "Optimized Feature Generation for Tabular Data via LLMs with Decision Tree Reasoning", *In Proceedings of the 38th Conference on Neural Information Processing Systems (NeurIPS 2024)*. https://doi.org/10.52202/079017-2932

Patil, D. (2024) "Explainable Artificial Intelligence (XAI) for industry applications: Enhancing transparency, trust, and informed decision-making in business operation", *Hurix Digital*. https://www.researchgate.net/publication/385881486

Penttala, J. (2025) "Performance of LLM models with numerical data", Master's Thesis, *University of Helsinki*. https://helda.helsinki.fi/items/aaa2fe24-fbcd-4cc4-a65d-912746098b92

Petukhova, A., Matos-Carvalho, J. P. and Fachada, N. (2024) "Text Clustering with Large Language Model Embeddings", *International Journal of Cognitive Computing in Engineering*. https://doi.org/10.1016/j.ijcce.2024.11.004

Salminen, J., Jung, S.-G., Chowdury, S. and Jansen, B. J. (2026) "AI-generated personas: Representing user needs with generative AI models", *International Journal of Human-Computer Studies*. https://doi.org/10.1016/j.ijhcs.2026.103745

Shutaywi, M. and Kachouie, N.N. (2021). "Silhouette analysis for performance evaluation in machine learning with applications to clustering." *Entropy*, 23(6), 759. https://doi.org/10.3390/e23060759

Sidek, Z. and Ahmad, S. S. S. (2026) "Exploring word embeddings and clustering algorithms for user reviews", *Indonesian Journal of Electrical Engineering and Computer Science*, Vol. 41, No. 3, pp. 1017-1024. https://doi.org/10.11591/ijeecs.v41.i3.pp1017-1024

Simpson, C., Tabatsky, E., Rahil, Z., Eddins, D. J., Tkachev, S., Georgescauld, F., Papalegis, D., Culka, M., Levy, T., Gregoretti, I., Meehan, C., Schiller, C., Bestak, K., Schapiro, D., Chernyshev, A., Walther, G., Ghosn, E. E. B. and Orlova, D. (2025) "Lifting the curse from high-dimensional data: Automated projection pursuit clustering for a variety of biological data modalities", *GigaScience*, Vol. 14, pp. 1-20. https://doi.org/10.1093/gigascience/giaf052

Smutek, T., Sikora, J., Bogacki, S., Rutkowski, M. and Woźniak, D. (2024) "Use of Autoencoder and One-Hot Encoding for Customer Segmentation", *European Research Studies Journal*, Vol. XXVII, Special Issue 2, pp. 72-82. https://doi.org/10.35808/ersj/3388

Sundqvist, M., Chiquet, J. and Rigaill, G. (2023). "Adjusting the adjusted Rand Index: A multinomial story". *Computational Statistics*, 38(1), 327–347. https://doi.org/10.1007/s00180-022-01230-7

Syahra, Y., Fadlil, A. and Yuliansyah, H. (2025). "Customer segmentation using RFM and K-Means clustering to support CRM in retail industry". *Sinkron: Jurnal dan Penelitian Teknik Informatika*, 9(3), 1120–1131. https://doi.org/10.33395/sinkron.v9i3.14907

Wang, P. and Ding, H. (2024) "The rationality of explanation or human capacity? Understanding the impact of explainable artificial intelligence on human-AI trust and decision performance", *Information Processing and Management*, Vol. 61, No. 4, Art. 103732. https://doi.org/10.1016/j.ipm.2024.103732

Woltmann, L., Thiele, M. and Lehner, W. (2018) "Modeling Customers and Products with Word Embeddings from Receipt Data", *IDEAS 2018: 22nd International Database Engineering & Applications Symposium*, pp. 246-252. https://doi.org/10.1145/3216122.3229860

Perturbation-First Robustness Evaluation of Automated Short Answer Grading

Ferdinando Sasso, Andrea De Mauro
LUISS Guido Carli University, Rome, Italy

Abstract

Automated Short Answer Grading (ASAG) is increasingly deployed in high-stakes educational contexts, yet the dominant evaluation paradigm focuses primarily on agreement with human raters, overlooking robustness as a complementary quality dimension. Robustness, defined as the consistency of grading behaviour under controlled, meaning-aware input transformations, bears directly on construct validity, the degree to which an assessment measures what it claims to measure, and conditions the trustworthiness of automated decisions in education.

We propose a perturbation-first evaluation framework that assesses robustness across three validity-theoretic dimensions: invariance to construct-irrelevant variation, sensitivity to meaning-altering changes, and resistance to adversarial gaming. The framework combines rule-based perturbation generators, two-gate quality validation, and self-referential metrics that compare the grader against its own output rather than against gold labels, isolating robustness from accuracy. A dual-protocol design contrasts within-question evaluation, the optimistic scenario typically reported in the literature, with cross-question evaluation, the realistic deployment scenario, capturing how trained graders degrade under distribution shift.

We apply the framework to a benchmark ASAG dataset of science short answers and compare a trained machine-learning baseline with zero-shot Large Language Model (LLM) graders evaluated at two prompting levels, with and without access to a reference answer. Three findings emerge. First, trained graders exhibit substantial robustness degradation under realistic cross-question evaluation, suggesting that standard within-question benchmarks overestimate operational reliability. Second, zero-shot LLM graders exhibit stronger robustness on invariance and sensitivity dimensions while remaining comparatively less vulnerable to adversarial manipulation, although residual sensitivity failures display a stable hierarchy across perturbation types. Finally, providing reference answers introduces a subtle trade-off, with improved consistency on meaning-preserving variations but increased susceptibility to keyword-based gaming, which is invisible to single-metric evaluation.

We argue that multi-dimensional robustness evaluation should complement accuracy-based metrics in ASAG systems intended for high-stakes educational use, and we release an open-source framework that operationalises the approach and is freely available to the research community.

Keywords: Automated Short Answer Grading; Robustness Evaluation; Perturbation Testing; Large Language Models; Construct Validity

Paper type: Academic Research Paper

1 Introduction

Automated Short Answer Grading (ASAG) systems assign grades to free-text student responses by comparison with reference answers. The dominant evaluation paradigm, which relies on accuracy or Quadratic Weighted Kappa against gold labels, reveals an important blind spot: a grader may agree with human raters on clean test data yet behave erratically when inputs vary in superficial, meaning-preserving ways. Such inconsistency violates construct validity and undermines the pedagogical purpose of assessment (Messick, 1989; Williamson et al., 2012). Operational ASAG systems are also vulnerable to simple adversarial inputs (Ding et al., 2020; Kumar et al., 2020; Filighera et al., 2024), yet such vulnerabilities remain largely absent from standard benchmarking.

We argue that robustness should be evaluated alongside accuracy as a first-class quality dimension and propose a perturbation-first framework built on three families targeting distinct validity threats: invariance (construct-irrelevant variance), sensitivity (construct underrepresentation), and gaming (adversarial manipulation). Four metrics compare the grader against itself rather than against gold labels, isolating robustness from accuracy, and a dual-protocol design (Leave-One-Question-Out vs. within-question 80/20 split) quantifies how much consistency a trained grader loses beyond its training distribution.

We apply this framework to the SemEval 2013 Beetle dataset (5,199 answers, 42 questions), evaluating a hybrid ML baseline and a Large Language Model (LLM) grader (GPT-5.4 mini) at two prompting levels. The contributions of this work are threefold: a perturbation-first evaluation framework grounded in validity theory; a dual-protocol design that exposes the extent to which within-question evaluations overestimate operational

reliability; and a cross-paradigm comparison that documents a structured sensitivity-blindness gradient and an unexpected gaming trade-off introduced by reference-answer prompting.

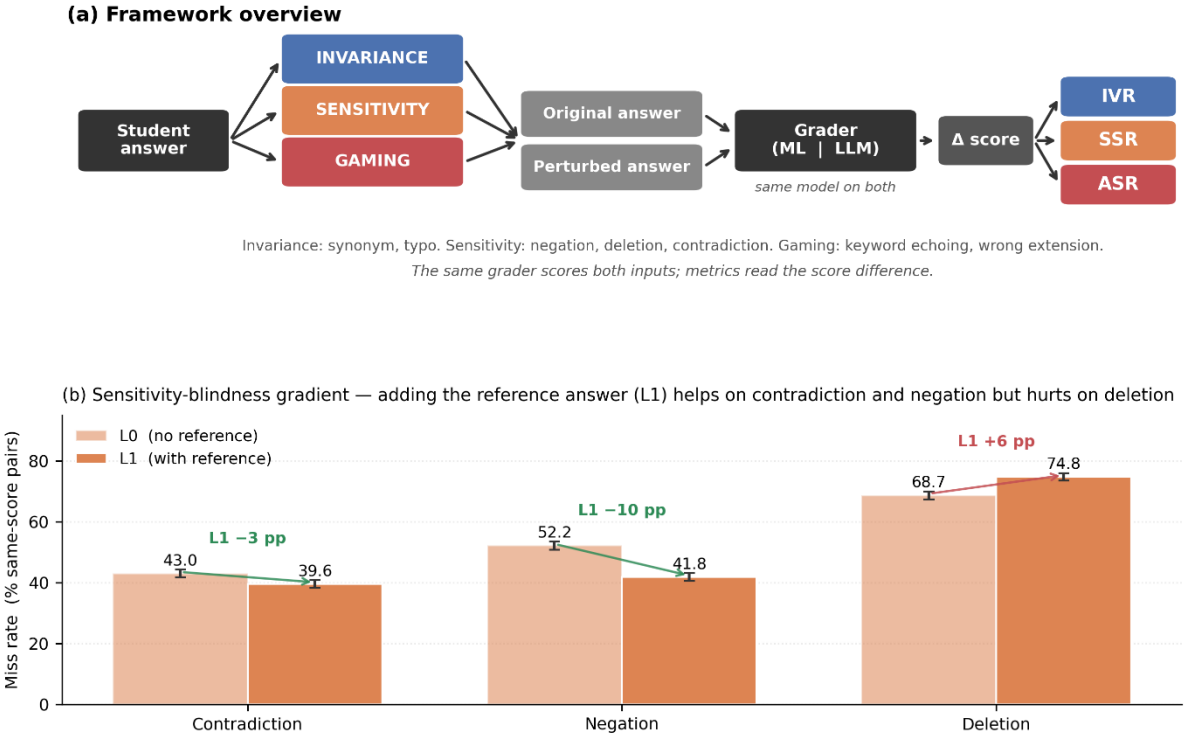


Figure 1: Framework overview and headline finding. (a) Pipeline: seven generators across three perturbation families produce paired (original, perturbed) inputs to the grader; three self-referential metrics — Invariance Violation Rate (IVR), Sensitivity Success Rate (SSR), and Adversarial Success Rate (ASR) — compare it against itself. (b) Sensitivity-blindness gradient in GPT-5.4 mini with 95% bootstrap confidence intervals. Reference prompting (L1) reduces missing on negation and contradiction but increases it on deletion: the L1 paradox.

2 Theoretical Background

The ASAG literature spans nearly two decades of methodological progress, from lexical and dependency-based approaches (Mohler et al., 2011; Burrows et al., 2015) to transformer-based models (Sung et al., 2019; Condor et al., 2021). Across this trajectory, evaluation has predominantly reported agreement with human raters via accuracy, Quadratic Weighted Kappa, or correlation, while the gap between in-distribution agreement scores and operational reliability has long been a concern of the educational measurement community (Heilman and Madnani, 2015; Williamson et al., 2012; Deane, 2013).

The limits of agreement-based evaluation are best understood through construct validity, that is, the degree to which an assessment measures what it claims to measure (Messick, 1989; Kane, 2006). Messick's framework identifies two principal threats: *construct-irrelevant variance* (scores influenced by factors unrelated to the construct, e.g. a grader penalising a correct answer for a typo) and *construct underrepresentation* (the assessment fails to capture meaningful aspects of the construct, e.g. a grader missing a negation that inverts meaning). To these, the literature on automated scoring has added a third threat specific to algorithmic systems, *adversarial gaming*, in which surface manipulations such as keyword stuffing artificially inflate scores (Williamson et al., 2012; Ferrara and Qunbar, 2022).

Perturbation-based evaluation is a standard methodology for assessing NLP model robustness. Ribeiro et al. (2020) introduced CheckList, a behavioural testing framework that systematically probes model capabilities through invariance, directional, and minimum-functionality tests. The general adversarial-NLP literature provides a wide toolkit of attack strategies, surveyed in Wang et al. (2022).

Within automated scoring, adversarial vulnerabilities have been documented but not systematically addressed. Ding et al. (2020) showed that content scoring systems could be fooled by random character strings, Kumar et al. (2020) attacked essay scoring with simple perturbations, and Filighera et al. (2024) extended this work to ASAG via adversarial adjectives and adverbs. Each focuses on a single attack family, and none proposes a unified, multi-dimensional framework anchored to validity theory.

The emergence of large language models has introduced new scoring paradigms that range from zero-shot or few-shot prompting (Mizumoto and Eguchi, 2023; Naismith et al., 2023) to lightweight fine-tuning (Latif and Zhai, 2024). These studies report encouraging accuracy figures but do not characterise robustness. Two questions are pressing: whether the cross-question robustness drop documented for trained models also affects zero-shot LLM graders, and whether prompting interventions (e.g. supplying a reference answer) interact uniformly with all validity dimensions or produce trade-offs invisible to accuracy alone.

3 Methodology

The evaluation framework follows a unidirectional pipeline (Figure 1a): data loading, perturbation generation, grading, metric computation, and cross-protocol comparison. All data flows through validated immutable models (questions, answers, perturbations, grade results) to ensure reproducibility. The framework is implemented in Python and released as an open-source pip-installable package (asag-perturbation) with a public API and command-line interface; source code, evaluation scripts, and the analyses reported in this paper are available at https://github.com/NNSS666/asag-perturbation-framework.

3.1 Perturbation Generation and Quality Validation

Seven rule-based generators produce up to ten perturbation variants per student answer, organised into three families.

Invariance family. Synonym substitution replaces content words with WordNet synonyms, producing up to two variants per answer (selected deterministically for reproducibility); typo insertion introduces a single character-level modification (swap, deletion, or duplication) into one content word.

Sensitivity family. Negation insertion adds "not" after auxiliary verbs or prepends "It is not true that" as a fallback. Key concept deletion removes one domain-relevant content word selected via seeded random sampling. Semantic contradiction replaces domain terms with curated antonyms (e.g., "open" → "closed", "series" → "parallel") drawn from a 40-pair dictionary covering physics and biology vocabulary.

Gaming family. Rubric keyword echoing appends reference-answer keywords absent from the student response, simulating keyword stuffing; fluent wrong extension appends a confident but factually incorrect domain statement from a curated pool of plausible-sounding misconceptions.

Illustrative examples for each perturbation type are reported below, with the perturbation marked in italics:

- *Synonym (invariance):* "the *bulb* lights up" → "the *lamp* lights up"
- *Negation (sensitivity):* "the terminals are connected" → "the terminals are *not* connected"
- *Deletion (sensitivity):* "the battery is in *a closed path*" → "the battery is in *a path*"
- *Contradiction (sensitivity):* "switches connected in *series*" → "switches connected in *parallel*"
- *Gaming:* "the bulb stays on" → "the bulb stays on *because voltage flows backwards*"

Invariance perturbations undergo two validation gates before acceptance. Gate 1 applies only to synonym substitution: candidate texts must achieve cosine similarity of at least 0.85 with the original, measured via sentence-BERT embeddings (all-MiniLM-L6-v2). Gate 2 applies to all invariance types: candidates introducing new negation markers or antonyms not present in the original are rejected. Rejected candidates are not regenerated, and the rejection rate is reported as a research result. The gates encode the *perturbation-level* ground truth (whether a candidate variant preserves meaning as intended), as distinguished from the *answer-level* gold labels supplied by the dataset (Section 4), which encode whether the original student answer is correct.

3.2 Robustness Metrics

All metrics compare the grader's score on the original answer against its score on the perturbed answer, isolating robustness from accuracy.

- *IVR_flip* (Invariance Violation Rate, binary): the proportion of invariance pairs where the score changed at all; lower is better.
- *IVR_absdelta* (Invariance Violation Rate, continuous): the mean absolute score difference across invariance pairs, capturing violation magnitude.
- *SSR_directional* (Sensitivity Success Rate): the proportion of sensitivity pairs where the perturbed score strictly decreased; higher is better, and no-change counts as a failure to detect the modification.
- *ASR_thresholded* (Adversarial Success Rate): the proportion of gaming pairs where the score crossed from below the passing threshold (0.5) to at or above it. Already-passing answers that increase further are excluded.

The four metrics are reported on their canonical scales rather than aggregated into a composite index, which preserves continuity with the validity-theoretic constructs each metric targets and avoids the arbitrariness of weighting heterogeneous dimensions whose practical importance varies by deployment context. Joint reading across the three dimensions, illustrated in Section 5.1, surfaces the trade-offs that single-metric evaluation hides.

3.3 Dual-Protocol Evaluation

The dual-protocol design targets trained graders and quantifies how much robustness they lose on unseen questions. Protocol A (Leave-One-Question-Out) implements cross-question evaluation: for each of 42 questions the grader is trained on answers from all other questions and tested on the held-out one, with a leakage diagnostic verifying that no held-out question text appears in training. This is the realistic deployment scenario. Protocol B (within-question 80/20) implements in-distribution evaluation: for each question independently, answers are split 80/20 with stratified sampling on gold labels and the grader is trained and tested on the same question. This is the optimistic scenario typically reported in the literature.

The *robustness drop* is the difference between A and B aggregate metrics, macro-averaged across questions. Zero-shot LLM graders receive an identical prompt regardless of split, so their A and B results are equivalent up to sampling noise, and we evaluate them on absolute metrics and compare against the trained baseline's Protocol A, the most demanding condition.

4 Experimental Setup

We use the SemEval 2013 Task 7 Beetle dataset (Dzikovska et al., 2013), comprising 5,199 student answers to 42 science questions about electrical circuits. Answers are labelled on a five-way scale (correct, partially_correct_incomplete, contradictory, non_domain, irrelevant) and normalised to a continuous [0, 1] scale (correct = 1.0, partial = 0.5, others = 0.0) following the ASAG2024 benchmark convention. These dataset-supplied labels constitute the *answer-level* ground truth used to interpret robustness results; the *perturbation-level* ground truth is enforced by the two-gate validation of Section 3.1 and by the rule-based construction of each perturbation type. The robustness metrics themselves do not consult either ground truth, since they compare the grader against itself.

4.1 Graders

Hybrid ML baseline. A logistic regression classifier trained on 388-dimensional feature vectors: four handcrafted linguistic features (lexical overlap, length ratio, negation flag, reference-token recall) concatenated with 384-dimensional sentence-BERT embeddings (all-MiniLM-L6-v2; Reimers and Gurevych, 2019). Class weights are balanced to handle label imbalance. This configuration is broadly representative of strong feature-based ASAG baselines (Sung et al., 2019; Condor et al., 2021).

LLM grader. Zero-shot prompting of GPT-5.4 mini at two information levels. Level 0 supplies the question and the student answer only; Level 1 additionally supplies the reference answer and instructs the model to compare the student response against it. Temperature is set to 0.0 for near-deterministic output, with a fixed seed for reproducibility, and the model outputs a JSON object with a label and a confidence score mapped to the same [0, 1] scale as the gold labels. Because the model operates in a zero-shot regime, the dual-protocol distinction does not alter its input, and it is evaluated once against the HybridGrader's Protocol A.

4.2 Perturbation Statistics

The perturbation engine generated 41,444 perturbations across all seven types from 5,199 source answers, averaging 8.0 perturbations per answer (theoretical maximum 10). Gate 1 rejected 40.3% of synonym substitution candidates, indicating that nearly half of WordNet synonym replacements produce semantically divergent texts in the science domain; Gate 2 provided additional filtering for meaning-inverting substitutions. The rejection rate is reported as a research result, since it quantifies the inherent difficulty of producing valid invariance perturbations in technical domains and motivates the two-gate design.

5 Results

Results are organised in four blocks: a dual-protocol analysis of the trained HybridGrader; a cross-paradigm comparison with the zero-shot LLM, including reference-answer availability (Section 5.1); score-distribution

artefacts and the floor effect (Section 5.2); and the structured sensitivity-blindness gradient (Section 5.3). We begin with the dual-protocol analysis.

Table 1 reports the HybridGrader's aggregate robustness metrics under both evaluation protocols.

Table 1: HybridGrader robustness metrics by evaluation protocol.

Metric	Protocol B (in-distribution)	Protocol A (cross-question)	Drop (A – B)
IVR_flip (↓)	0.179	0.337	+0.158
IVR_absdelta (↓)	0.125	0.242	+0.118
SSR_directional (↑)	0.120	0.140	+0.020
ASR_thresholded (↓)	0.115	0.188	+0.074

Invariance violations nearly double from Protocol B to A (IVR_flip 0.179 → 0.337): the grader changes its score on meaning-preserving perturbations 33.7% of the time on unseen questions, versus 17.9% on familiar ones. Gaming vulnerability follows the same pattern (ASR 0.115 → 0.188), with adversarial inputs 64% more effective on novel questions. Sensitivity detection remains low under both protocols (SSR ≈ 0.13), indicating that the HybridGrader's failure to detect meaning-altering perturbations is a structural limitation of feature-based scoring rather than a distribution-shift effect. The +15.8 pp IVR_flip and +7.4 pp ASR drops quantify the extent to which within-question evaluations overestimate operational robustness.

5.1 Cross-Paradigm Comparison: Trained ML vs. Zero-Shot LLM

Table 2 reports the zero-shot LLM's absolute robustness metrics, compared against the HybridGrader's Protocol A results.

Table 2: Cross-paradigm robustness comparison (HybridGrader at Protocol A; LLM zero-shot, single-run).

Metric	HybridGrader (A)	GPT-5.4 mini L0	GPT-5.4 mini L1
IVR_flip (↓)	0.337	0.243	0.134
IVR_absdelta (↓)	0.242	0.126	0.071
SSR_directional (↑)	0.140	0.420	0.461
ASR_thresholded (↓)	0.188	0.067	0.118

Four observations can be drawn from the comparison. *First, the LLM is more robust to surface variation.* GPT-5.4 mini Level 0 achieves IVR_flip of 0.243 versus 0.337 for the HybridGrader under Protocol A (a 28% relative reduction); with the reference answer available (Level 1) invariance violations fall to 0.134, a 60% reduction over the baseline. The reference appears to act as a semantic anchor against which meaning-preserving variations are recognised.

Second, the LLM detects meaning-altering perturbations more effectively. SSR_directional rises from 0.140 for the HybridGrader to 0.420 at Level 0 and 0.461 at Level 1, an approximately threefold improvement. The HybridGrader misses 86% of negations, deletions and contradictions, whereas the LLM at Level 1 misses 54%. Adding the reference answer yields only a modest additional gain (+4 pp), indicating that the sensitivity advantage arises predominantly from semantic processing rather than from the comparison target.

Third, reference-answer availability introduces an unexpected gaming trade-off. Providing the reference answer (Level 1) increases gaming susceptibility by 76% relative to Level 0 (ASR 0.118 vs 0.067). The prompt instructs the model to compare the student answer against the reference, so any token overlap becomes evidence of correctness, and when a gaming perturbation injects reference keywords into a wrong response the model registers them as partial understanding. The reference functions as a checklist from which an adversarial student can copy, whereas Level 0, lacking this anchor, must judge the answer on its overall coherence and is correspondingly harder to game.

Fourth, multi-metric evaluation reveals trade-offs invisible to any single metric. On invariance alone, Level 1 dominates Level 0; on gaming alone, Level 0 dominates Level 1; on sensitivity alone, the two are comparable. The same point applies to the trained baseline: under Protocol B its IVR_flip of 0.179 looks reasonable, yet SSR_directional of 0.120 reveals that this stability partly reflects indiscrimination, with the grader rarely changing its score even when it should.

5.2 Score Distribution Effects

To rule out the possibility that the lower IVR is an artefact of score compression, we examined the LLM's label distribution. At Level 0, GPT-5.4 mini exhibits a conservative scoring pattern: 54.3% of answers receive the partially_correct_incomplete label (score 0.5), against 23.1% in the gold distribution. Compression toward the midpoint mechanically reduces IVR, since scores at 0.5 have less room to flip. At Level 1, the distribution becomes substantially more discriminative, closer to the gold distribution, yet IVR_flip continues to fall (0.243 → 0.134). This rules out the score-compression explanation, and the invariance robustness observed reflects genuine semantic stability rather than a statistical artefact.

A symmetric artefact affects SSR_directional: answers already scored 0.0 cannot decrease further under sensitivity perturbations, mechanically producing SSR = 0% on those pairs regardless of grader quality. We term this the floor effect. Table 3 reports SSR with and without floor-affected pairs.

Table 3: SSR floor-effect analysis — overall vs. restricted to answers with original score > 0.

Configuration	Pairs	Floor pairs (orig=0)	SSR overall	SSR non-floor (orig>0)	Delta
GPT-5.4 mini L0	13,155	3,012 (22.9%)	0.423	0.549	+12.6 pp
GPT-5.4 mini L1	12,309	4,219 (34.3%)	0.480	0.731	+25.0 pp

The floor effect is larger for Level 1 because the reference answer makes the grader more discriminative, increasing the proportion of untestable pairs. After correction Level 1 detects 73.1% of meaning-altering perturbations on testable pairs, against 48.0% uncorrected.

5.3 Sensitivity Blindness Gradient

Disaggregating SSR by perturbation type reveals that the LLM's sensitivity failures follow a stable hierarchy (Figure 1b, Table 4): the grader is blindest to deletions, intermediate on negations, and least blind on outright contradictions, and the pattern holds for both prompting levels.

Table 4: Per-type miss rates (% same-score pairs) with 95% bootstrap CI (10k resamples).

Grader	Deletion	Negation	Contradiction
L0	68.7 [67.4, 70.0]	52.2 [50.8, 53.5]	43.0 [41.7, 44.3]
L1	74.8 [73.6, 76.0]	41.8 [40.5, 43.2]	39.6 [38.3, 40.9]

Pairwise McNemar tests confirm that the gradient is significant (all six within-grader comparisons $p < 0.001$). A further side-effect emerges between graders: adding the reference answer (L1) reduces the miss rate on negation (−10 pp) and contradiction (−3 pp) but *increases* it on deletion (+6 pp), a paradox confirmed by paired McNemar on the same answer set ($\chi^2 = 51$, $p < 0.0001$). Floor adjustment (Section 5.2) lowers all three rates uniformly without changing the ordering, and even at the bottom (~40% on contradiction in L1) the failure rate constitutes a substantive deployment risk.

A manual review of 30 random missed deletions characterises the mechanism. Of 17 cases involving deletion of a domain concept (terminal, battery, circuit, voltage, open/closed), 16 materially degraded or inverted answer meaning, while connective deletions ("because", "between") were correctly tolerated as benign. The grader systematically fails on content-bearing words. A representative inverted case is "the battery is in *a closed path*" → "the battery is in *a path*" (gold = correct, score 1.0 → 1.0): the closed/open distinction is the central conceptual axis of Beetle, and the Level 1 grader treats the two versions as equivalent.

6 Discussion, Implications, and Limitations

In this section we discuss the results of the study, distill implications for practitioners and researchers, and outline opportunities for further investigation.

The most consequential methodological finding is that within-question evaluations systematically overestimate the robustness of trained ASAG graders. The HybridGrader's invariance violation rate nearly doubles between the in-distribution and cross-question protocols, and its gaming susceptibility rises by roughly two-thirds, effects large enough to materially affect deployment. Published ASAG evaluations should therefore include cross-question results alongside in-distribution ones, and report the gap explicitly as a robustness drop.

The advantage that the zero-shot LLM exhibits on invariance and sensitivity arises from a form of semantic processing that feature-based scoring lacks. The HybridGrader relies on lexical overlap that shifts with surface

form, whereas the LLM processes the answer as language: a negation, deletion or antonym leaves surface features intact while shifting meaning, defeating feature-based scoring yet remaining at least partially visible to a semantic processor. The residual weaknesses in Section 5.3 refine rather than undermine this conclusion, showing that semantic processing handles meaning-preserving variation more reliably than meaning-altering omission.

Adding the reference answer is an intervention practitioners commonly expect to improve grading uniformly. It does so for invariance and sensitivity, but for gaming it nearly doubles vulnerability. The mechanism generalises beyond the particular prompt employed here: when the model is instructed to compare the student answer against a reference, any token overlap becomes evidence of correctness, which an adversarial student can exploit. The same checklist mechanism plausibly explains the deletion paradox of Section 5.3: when a domain concept is removed, the reference has no token to compare against while surrounding tokens still match, so the grader reads continued overlap as continued correctness and tolerates the omission. Reference-anchored prompting therefore fails on the same axis on which it succeeds, since token-level matching detects substitutions and negations but cannot detect what is no longer present. Prompt engineering and coherence-first instructions are plausible mitigations, yet such trade-offs remain invisible to single-metric evaluation.

As an outcome of the study, several implications can be drawn for practitioners and researchers. Practitioners should conduct perturbation evaluation across the three validity dimensions before placing a grader in consequential use, and should compare alternative prompting strategies on all dimensions rather than on accuracy alone, since gains in agreement with human raters can mask increased gaming exposure. Cross-question performance should be treated as the realistic baseline against which deployment decisions are made. From a research standpoint, the present findings position robustness alongside fairness as a precondition for responsible automated scoring (Loukina et al., 2019), and motivate further investigation of prompt designs capable of mitigating the trade-offs documented here.

We acknowledge several limitations that could be object of further research. First, the empirical evidence is restricted to a single dataset (Beetle) covering one science domain, and replication on additional ASAG corpora is necessary to establish the generality of the patterns reported. Second, the corpus lacks structured per-question rubrics, so the contrast between Level 0 and Level 1 prompting serves as a proxy for how additional grading information shapes robustness. Third, the LLM evaluation relies on a single model (GPT-5.4 mini), and cross-vendor as well as frontier-model comparisons are deferred to follow-up work. Finally, perturbations are produced by rule-based generators, and complementary human validation through inter-annotator agreement is a natural extension for the journal version.

7 Conclusions

We presented a perturbation-first evaluation framework for ASAG that quantifies robustness along three validity dimensions from Messick's taxonomy: invariance, sensitivity, and gaming. The framework combines seven rule-based generators, two-gate quality validation, four self-referential metrics, and a dual-protocol design that quantifies how trained graders degrade under cross-question generalisation.

Applied to the SemEval 2013 Beetle corpus, three findings stand out. The trained ML baseline exhibits a substantial robustness drop under cross-question evaluation (+15.8 pp IVR_flip; +7.4 pp ASR), confirming that within-question benchmarks overestimate operational reliability. The zero-shot LLM grader is markedly more robust on invariance and sensitivity, and less vulnerable to gaming. Adding the reference answer trades a 45% reduction in invariance violations for a 76% increase in gaming susceptibility, a subtle trade-off invisible to any single metric. We argue that robustness should be evaluated alongside accuracy as a first-class quality dimension for any ASAG system intended for high-stakes use, and that such evaluation should be multi-dimensional and grounded in validity theory.

8 Ethics Declaration

This study uses publicly available data from the SemEval 2013 Task 7 Beetle benchmark, which contains anonymised student responses released by the dataset authors for research purposes (Dzikovska et al., 2013). No primary data collection involving human participants was carried out for the present study, and no ethical clearance was required.

9 AI Declaration

Artificial intelligence tools were used in two ways. First, OpenAI's GPT-5.4 mini was the object of evaluation as one of the graders studied; its prompting configuration, decoding parameters, and access conditions are

documented in Section 4.1, and the resulting outputs are reported without modification. Second, general-purpose LLMs (Anthropic Claude and OpenAI GPT) were employed as writing assistants during manuscript preparation, specifically for grammar and style proof-reading, terminology consistency checks, and rephrasing suggestions on author-drafted text. All scientific content, research design, analyses, results, and interpretations are the authors' own. AI-generated wording was reviewed, edited, and retained only where consistent with the authors' intent and with the underlying empirical evidence.

References

Burrows, S., Gurevych, I. and Stein, B. (2015) 'The eras and trends of automatic short answer grading', *International Journal of Artificial Intelligence in Education*, 25(1), pp. 60–117.

Condor, A., Litster, M. and Pardos, Z. (2021) 'Automatic short answer grading with SBERT on out-of-sample questions', in *Proceedings of the 14th International Conference on Educational Data Mining (EDM 2021)*. International Educational Data Mining Society, pp. 376–382.

Deane, P. (2013) 'On the relation between automated essay scoring and modern views of the writing construct', *Assessing Writing*, 18(1), pp. 7–24.

Ding, Y., Riordan, B., Horbach, A., Cahill, A. and Zesch, T. (2020) 'Don't take "nswvtnvakgxpm" for an answer — The surprising vulnerability of automatic content scoring systems to adversarial input', in *Proceedings of the 28th International Conference on Computational Linguistics (COLING 2020)*. Barcelona, pp. 882–892.

Dzikovska, M., Nielsen, R., Brew, C., Leacock, C., Giampiccolo, D., Bentivogli, L., Clark, P., Dagan, I. and Dang, H.T. (2013) 'SemEval-2013 Task 7: The joint student response analysis and 8th recognizing textual entailment challenge', in *Proceedings of the 7th International Workshop on Semantic Evaluation (SemEval 2013)*. Association for Computational Linguistics, pp. 263–274.

Ferrara, S. and Qunbar, S. (2022) 'Validity arguments for AI-based automated scores: essay scoring as an illustration', *Journal of Educational Measurement*, 59(3), pp. 288–313.

Filighera, A., Ochs, S., Steuer, T. and Tregel, T. (2024) 'Cheating automatic short answer grading with the adversarial usage of adjectives and adverbs', *International Journal of Artificial Intelligence in Education*, 34(2), pp. 616–646.

Heilman, M. and Madnani, N. (2015) 'The impact of training data on automated short answer scoring performance', in *Proceedings of the 10th Workshop on Innovative Use of NLP for Building Educational Applications (BEA 2015)*. Denver: Association for Computational Linguistics, pp. 81–85.

Kane, M.T. (2006) 'Validation', in Brennan, R.L. (ed.) *Educational Measurement*. 4th edn. Westport, CT: American Council on Education/Praeger, pp. 17–64.

Kumar, Y., Bhatia, M., Kabra, A., Li, J.J., Jin, D. and Shah, R.R. (2020) 'Calling out bluff: Attacking the robustness of automatic scoring systems with simple adversarial testing', *arXiv preprint* arXiv:2007.06796.

Latif, E. and Zhai, X. (2024) 'Fine-tuning ChatGPT for automatic scoring', *Computers and Education: Artificial Intelligence*, 6, p. 100210.

Loukina, A., Madnani, N. and Zechner, K. (2019) 'The many dimensions of algorithmic fairness in educational applications', in *Proceedings of the 14th Workshop on Innovative Use of NLP for Building Educational Applications (BEA 2019)*. Florence: Association for Computational Linguistics, pp. 1–10.

Messick, S. (1989) 'Validity', in Linn, R.L. (ed.) *Educational Measurement*. 3rd edn. Washington, DC: American Council on Education/Macmillan, pp. 13–103.

Mizumoto, A. and Eguchi, M. (2023) 'Exploring the potential of using an AI language model for automated essay scoring', *Research Methods in Applied Linguistics*, 2(2), p. 100050.

Mohler, M., Bunescu, R. and Mihalcea, R. (2011) 'Learning to grade short answer questions using semantic similarity measures and dependency graph alignments', in *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics (ACL 2011)*. Portland, pp. 752–762.

Naismith, B., Mulcaire, P. and Burstein, J. (2023) 'Automated evaluation of written discourse coherence using GPT-4', in *Proceedings of the 18th Workshop on Innovative Use of NLP for Building Educational Applications (BEA 2023)*. Toronto, pp. 394–403.

Reimers, N. and Gurevych, I. (2019) 'Sentence-BERT: Sentence embeddings using Siamese BERT-networks', in *Proceedings of EMNLP-IJCNLP 2019*. Hong Kong: Association for Computational Linguistics, pp. 3982–3992.

Ribeiro, M.T., Wu, T., Guestrin, C. and Singh, S. (2020) 'Beyond accuracy: Behavioral testing of NLP models with CheckList', in *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL 2020)*. Association for Computational Linguistics, pp. 4902–4912.

Sung, C., Dhamecha, T.I. and Mukhi, N. (2019) 'Improving short answer grading using transformer-based pre-training', in *Proceedings of the 20th International Conference on Artificial Intelligence in Education (AIED 2019)*. Lecture Notes in Computer Science, vol. 11625. Cham: Springer, pp. 469–481.

Wang, X., Wang, H. and Yang, D. (2022) 'Measure and improve robustness in NLP models: A survey', in *Proceedings of NAACL 2022*. Seattle: Association for Computational Linguistics, pp. 4569–4586.

Williamson, D.M., Xi, X. and Breyer, F.J. (2012) 'A framework for evaluation and use of automated scoring', *Educational Measurement: Issues and Practice*, 31(1), pp. 2–13.

Artificial Intelligence Applications in Biogas Systems: A Systematic Literature Review

Vincenzo Fiengo, Renato Passaro, Ivana Quinto
University "Parthenope", Naples, Italy

Abstract

The increasing importance of renewable energy and circular economy approaches has made biogas an important technology in sustainable energy generation and waste management (Uddin and Wright, 2023). In this respect, the application of artificial intelligence (AI) and machine learning (ML) in improving the efficiency and management of anaerobic digesters has become increasingly popular (Ananda et al., 2025). While there has been considerable research activity in this field, a systematic understanding of its progress has not yet been achieved. This research aims to fill this gap by performing a systematic literature review (SLR) based on PRISMA criteria (Page et al., 2021). A total of 131 articles has been selected from Scopus and Web of Science databases for analysis during the period 2010–2025. This paper presents a comprehensive overview of the literature using a combination of bibliometric and thematic analysis. Findings indicate a significant increase in the number of publications in recent years, confirming the growing scientific and practical interest in the topic. Moreover, there is a dominance of technical or engineering studies with a multidisciplinary yet fragmented body of literature, highlighting the evolving nature of the research field. From the thematic analysis, three major groups of studies were identified, which include: i) strategic and sustainability-oriented applications, focusing on techno-economic and environmental assessment; ii) operational optimisation, where AI is used to improve process efficiency, prediction, and control; iii) advanced computational modelling, centred on the development of hybrid and data-driven algorithms. The findings highlight that AI and ML are primarily used to enhance predictive accuracy and operational performance, while their potential as decision-support tools at strategic and organisational levels remains underexplored. Accordingly, the study proposes a research agenda emphasising the need to integrate AI into decision-making processes, improve data-driven management practices, and develop more holistic and interdisciplinary approaches to biogas system management.

Keywords: Anaerobic digestion, Artificial intelligence, Biogas production, Circular economies, Digital technologies.

Paper type: Academic Research Paper

1 Introduction

Over the past decades, growing issues over climate change, environmental pollution, and rising energy demand have determined a global shift from fossil fuels toward renewable energy sources (Overland, Juraev and Vakulchuk, 2022). In this context, biogas has emerged as a valuable renewable energy source, produced from biomass and organic waste, where its development driven by the expansion of agricultural and waste management activities (Salehian, 2013). The core process to enable biogas production is Anaerobic Digestion (AD). AD is a biological process where microorganisms decompose organic matter in the absence of oxygen, producing biogas mainly composed of methane and carbon dioxide. The process takes place in closed reactors, and the generated biogas is used primarily for heat and electricity production. From a circular economy perspective, anaerobic digestion contributes to closing material and energy loops by converting organic waste into renewable energy and nutrient-rich digestate, thereby reducing waste disposal and fossil energy dependence (Uddin and Wright, 2023).

In recent years, AI and ML have been increasingly applied to improve process efficiency, support techno-economic analysis (TEA) (Dominic and Baidurah, 2025), and assess environmental impacts in energy and biogas systems (Ranieri et al.,2023). Considering the growing interesting, in recent years biogas production and AD have been addressed from different but complementary perspectives. On one hand, some studies have mainly focused on the technological and process-related aspects of biogas production (Uranga-Valencia et al., 2025), while on the other hand recent papers have investigated the integration of automation, AI and ML in biomass and bioenergy systems, highlighting the potential of data-driven methods for process monitoring, prediction, and optimisation (Ananda, Sujeeth and Archana, 2025). Despite the growing number of papers focusing on A /ML applications in biogas production and AD, to our knowledge, it is still lacking studies able to properly investigate these topics in a systematic and reproducible way. Put differently, a comprehensive and rigorous analysis of the research landscape is still missing, including structured analyses of publication trends, influential contributions, and thematic development over time. This is particularly relevant given the fragmented and

multidisciplinary nature of the field, where a quantitative mapping of the interconnections between biogas research and digital technologies is essential to identify structural patterns, emerging themes, and existing knowledge gaps. In line with this, the main aim of this paper is to provide a comprehensive and systematic review of the literature on AI and ML applications in biogas and AD systems, with a particular focus on their economic, environmental, and operational implications.

Therefore, the study aims to address the following research questions:

- *RQ1: How is the literature about the applications of AI and Machine Learning in biogas systems developing?*
- *RQ2: What are the main research themes emerging from the literature on AI and machine learning in biogas systems?*
- *RQ3: Which artificial intelligence and machine learning techniques are most applied and for which operational purposes (e.g., prediction, optimization, monitoring)?*
- *RQ4: What economic, environmental, and operational benefits are associated with the adoption of AI and machine learning in biogas plants?*
- *RQ5: What are the implications and future research directions related to the adoption of AI and machine learning in biogas systems?*

2 Methodology

This study performs a systematic literature review conducted in accordance with the PRISMA guidelines (Page et al., 2021). The adopted approach is consistent with established frameworks for conducting literature reviews in management and interdisciplinary research (Snyder, 2019), ensuring transparency, rigor, and reproducibility in the identification, selection, and analysis of the relevant studies.

Anaerobic digestion-based biogas systems are considered not only as technical systems but also as organizational entities, whose performance depends on technological, managerial, and operational factors. In this context, data-driven technologies can support decision-making, process optimization, monitoring, and sustainability evaluation. To collect the required data, two recognized databases have been used: 1. Scopus; and 2. Web of Science. The search timeframe spanned from January 1, 2010, to December 31, 2025, with the objective of compiling all articles published during that time. The necessary articles were collected through a keyword search in English, using a combination of Boolean operators. The keywords used were the following:

(TITLE-ABS-KEY (biodigester) OR TITLE-ABS-KEY ("biogas production technology") OR TITLE-ABS-KEY ("biogas production plant") OR TITLE-ABS-KEY ("biogas production firm")) AND (TITLE-ABS-KEY ("machine learning") OR TITLE-ABS-KEY ("artificial intelligence")).*

Within the framework of this research, specific inclusion criteria were defined, considering the following aspects: i) types of samples: studies focusing on biogas production through AD were included; ii) digital technologies: articles addressing the application of ML and/or AI to biogas production systems were considered; iii) language of articles: only articles written in English were included; iv) nature of articles: the review focused on peer-reviewed journal articles, including both original research and review papers, published in scientific journals and indexed in the selected databases; v) period of publication: only articles published between January 1, 2010, to December 31, 2025, were considered. Three reviewers independently conducted the data extraction process, evaluating all selected articles according to the predefined inclusion and exclusion criteria. To ensure consistency and reproducibility, a structured data extraction form was developed using Microsoft Excel. For each selected article, a structured data extraction process was conducted. The following information was collected: I) authors' names; ii) year of publication; iii) journal of publication; iv) geographic origin of the study; v) objective of the study; vi) type of biogas system analyzed; vii) AI/ML techniques applied; viii) purpose of the application; ix) input variables and output indicators; x) model performance metrics; xi) main findings. A descriptive and content analysis was then performed by computing certain variables (distributions per year, citation indexes, etc.) and by in-depth analysis of selected papers. Finally, a critical analysis and discussion of the results obtained was carried out. The descriptive analysis involved a comprehensive overview of the selected articles. To investigate relational networks such as co-authorship and bibliographic coupling the software VOSviewer was used. Bibliometric analysis was performed following recognized approaches (Donthu et al., 2021; Aria & Cuccurullo, 2017), while content analysis was conducted based on established qualitative frameworks (Krippendorff, 2018; Mayring, 2014).

3 Results

Articles were collected in a database that was interrogated through Vos Viewer to obtain bibliometrics and visualisations. The results of the SLR are organised and presented into the following sections:

1. Corpus overview and bibliometrics
   - 1.1 Temporal trend
   - 1.2 Papers per source
   - 1.3 Citations per years and top-cited papers
   - 1.4 Most influential authors and cited papers
   - 1.5 Subject area distribution
   - 1.6 Co-authorship map
   - 1.7 Common keywords
2. Content analysis of the sources

3.1 Corpus overview and bibliometrics

The final corpus consists of 131 articles, for which a bibliometric analysis was conducted to provide a general overview of the investigated topic.

The selection process is summarized in the PRISMA flow diagram. During the systematic review, a total of 218 documents were identified in the Scopus and Web of Science databases. Subsequently, 25 duplicate records were removed. After screening the abstracts of the remaining 193 papers, 45 articles were excluded. A further 17 papers were deemed not relevant following full-text assessment. Ultimately, 131 articles were included in the final sample.

The temporal distribution of publications highlights a clear increase in research activity over time. Between 2010 and 2019, the number of contributions remained limited, ranging from 0 to 3 publications per year. From 2020 onwards, a significant growth is observed, reaching a peak in 2025 with 37 publications. This trend indicates a rapidly growing interest in the field, particularly in recent years.

As reported in Table 1, the publications in the sample appear to be highly multidisciplinary and published in a variety of journals. Bioresource Technology is the main contributor with 11 papers, followed by energies for which 6 articles collected.

Table 1: Top 10 journal

Journal	N° papers
Bioresource Technology	11
Energies	6
Fuel	5
Journal Of Environmental Management	5
Science Of the Total Environment	5
Biomass & Bioenergy	4
Journal Of Cleaner Production	4
Waste Management	4
Biomass Conversion And Biorefinery	3
Chemosphere	3

Reporting the number of citations per year, Table 2 shows that the highest number of citations is observed in 2023. A marked increase in citations is evident during the period 2020–2024, confirming a growing scientific impact of the research topic.

Table 2: Trend of citations and number of articles per year

Year	Number of articles	Citations
2010	1	196
2011	0	0
2012	0	0
2013	0	0
2014	0	0
2015	1	66
2016	2	143
2017	3	76
2018	3	196

Year	Number of articles	Citations
2019	3	164
2020	6	618
2021	9	541
2022	13	664
2023	24	734
2024	26	442
2025	37	149

The top 10 most cited articles are summarized in Table 3. Cruz et al. (2022) is the most cited paper with 211 citations and the highest CPY. Bioresource Technology plays a key role, with several contributions represented in the sample. Most of top-cited papers focus on the application of machine learning techniques, such as artificial neural networks, genetic algorithms, and hybrid models, for predicting and optimizing biogas production processes. This last finding clearly suggest that the research field is evolving towards the integration of machine learning methods for improving the performance and understanding of anaerobic digestion systems.

Table 3: Most influential authors and cited papers

N°	Authors	Title	Year	Journals	Citations	CPY
1	Cruz, IA; Chuenchart, W; Long, F; Surendra, KC; Andrade, LRS; Bilal, M; Liu, H; Figueiredo, RT; Khanal, SK; Ferreira, LFR	Application of machine learning in anaerobic digestion: Perspectives and challenges	2022	BIORESOURCE TECHNOLOGY	211	70.33
2	Wang L.; Long F.; Liao W.; Liu H.	Prediction of anaerobic digestion performance and identification of critical operational parameters using machine learning algorithms	2020	Bioresource Technology	196	39.20
3	Abu Qdais, H; Hani, KB; Shatnawi, N	Modeling and optimization of biogas production from a waste digester using artificial neural network and genetic algorithm	2010	RESOURCES CONSERVATION AND RECYCLING	196	13.07
4	Kana, EBG; Oloke, JK; Lateef, A; Adesiyan, MO	Modeling and optimization of biogas production on saw dust and other co-substrates using Artificial Neural network and Genetic Algorithm	2012	RENEWABLE ENERGY	176	13.54
5	De Clercq, D; Wen, ZG; Fei, F; Caicedo, L; Yuan, K; Shang, RX	Interpretable machine learning for predicting biomethane production in industrial-scale anaerobic co-digestion	2020	SCIENCE OF THE TOTAL ENVIRONMENT	169	33.80
6	Almomani, F	Prediction of biogas production from chemically treated co-digested agricultural waste using artificial neural network	2020	FUEL	144	28.80
7	Najafi, B; Ardabili, SF	Application of ANFIS, ANN, and logistic methods in estimating biogas production from spent mushroom compost (SMC)	2018	RESOURCES CONSERVATION AND RECYCLING	140	20.00
8	Neto, JG; Ozorio, LV; de Abreu, TCC; dos Santos, BF; Pradelle, F	Modeling of biogas production from food, fruits and vegetables wastes using artificial neural network (ANN)	2021	FUEL	129	32.25

N°	Authors	Title	Year	Journals	Citations	CPY
9	Beltramo, T; Ranzan, C; Hinrichs, J; Hitzmann, B	Artificial neural network prediction of the biogas flow rate optimised with an ant colony algorithm	2016	BIOSYSTEMS ENGINEERING	102	11.33
10	Tufaner, F; Demirci, Y	Prediction of biogas production rate from anaerobic hybrid reactor by artificial neural network and nonlinear regressions models	2020	CLEAN TECHNOLOGIES AND ENVIRONMENTAL POLICY	99	19.80

Using functionalities provided by the SCImago Journal Rank (SJR) platform, fifty-three journal subject areas have been identified. Two main aspects emerge from the subject distribution. First, the topic is primarily covered by technical and engineering disciplines such as energy, environmental science, and chemistry, while socio-economic areas play a more marginal role. Second, the presence of many subject categories indicates a high degree of fragmentation, suggesting that the research field is still evolving. Third, the publication distribution pattern across many different subjects demonstrates the multidisciplinary character of the field. The multidisciplinary composition not only reflects the multifaceted of biogas systems that need to be optimized technologically and evaluated as whole systems, but it also suggests poor integration of technical and socio-economic considerations.

Fig.1 presents the co-authorship map obtained using the software VOSviewer, whereby nodes are the authors. 429 out of 753 meet the thresholds: a minimum number of one document for an author, and a minimum number of one citation for an author. From the network visualization, several distinct clusters are recognizable, each corresponding to research groups or established collaborations between scholars. In each cluster there is a dense connection, indicating strong scientific cooperation between authors belonging to the same group. Despite the lack of cooperation among the various groups, some authors such as Lim Jun Wei and Chan Yi Jing act as key connectors within the network, linking different clusters and facilitating knowledge exchange across research groups.

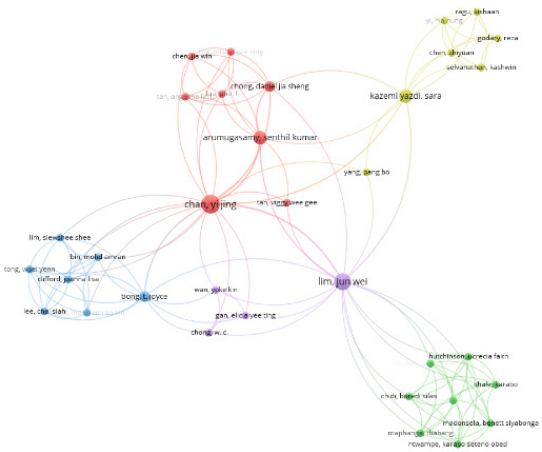


Fig.1: Co-authorship map

Using the same approach, a keyword co-occurrence analysis was conducted with VOSviewer to explore the main research themes in the field (Figure 2). 1600 keywords were obtained that met of the threshold of a minimum number of occurrences set equal to five. The map highlights a highly interconnected structure, with some central concepts acting as the hub of the entire network. “biogas” and “machine learning” emerge as the most relevant nodes, suggesting that the literature analysed focuses mainly on the application of ML techniques in the field of biogas production and optimization.

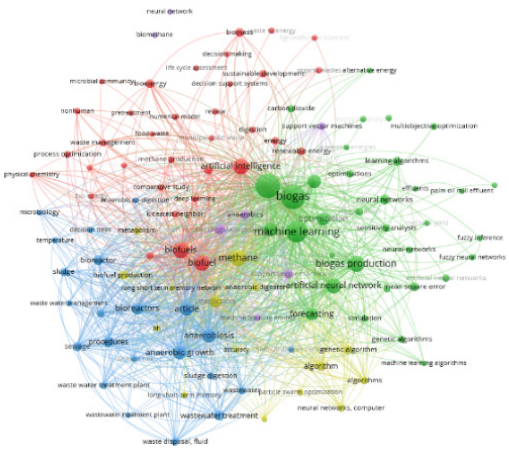


Figure 2: Co-occurrence keyword map

3.2 Content analysis

Content analysis recognises the importance of understanding the context in which an investigated item appears (Bryman, 2004). Based on the content analysis, the selected 131 papers are grouped in three main area, which we defined as “clusters”.

The first cluster focuses on strategic, environmental, and economic impacts of the adoption of AI in biodigester. Investigating how AI supports sustainability assessment and TEA, enabling scenario evaluation, environmental impact modelling, and performance benchmarking.

The second cluster represents the technical-operational core of this study. Here, the biodigester is conceptualised as a company that must forecast output, optimise yields, reduce operational uncertainty, and improve overall performance.

Finally, the third cluster focuses on the development and refinement of advanced algorithmic tools that enable predictive modelling and optimisation in biogas systems. The main contributions in this cluster include deal with optimisation strategies, development of hybrid modelling architectures and complex simulations and scenarios analysis to improve decision robustness.

In the following sections, the three clusters are analysed and discussed.

3.2.1. Strategic and Sustainability-Oriented AI Applications

Cluster 1 focuses on the strategic, environmental, and economic positioning of the biogas plant within complex territorial and energy systems. The cluster includes 27 contributions, almost equally divided between empirical studies and review papers.

About review papers, Heiker et al. (2021) provide a systematic overview of modelling approaches for biogas production, showing the complementarity between top-down and bottom-up models. Top-down models offer an aggregated perspective, used to analyse energy supply-demand dynamics, regional potential, and emissions, while bottom-up models provide a detailed representation of technical processes, allowing to capture the dynamic behaviour of anaerobic digestion and enabling operational optimisation. Similarly, Ananda et al. (2025) review the integration of automation and AI in biomass conversion systems, focusing on data-driven approaches for optimisation and decision support.

Among the key empirical studies, Dominic and Baidurah (2025) focus on techno-economic analysis, estimating key economic indicators and assess the feasibility of palm oil mill waste treatment systems. While Matheri et al. (2018) combine technical design considerations with decision-support methodologies for biodigester selection. But instead, Ouderji et al. (2023) provide a different prospective, integrating ML models with thermodynamic analysis and life cycle assessment, enabling the comparison of alternative energy configurations based on overall environmental performance and supporting decision-making processes through the identification of trade-offs between energy production, consumption, and environmental impact. Similarly, Selvanathan et al. (2023) apply ML techniques to predict methane emissions from palm oil mill effluent treatment systems, using real operational data to support environmental monitoring and emission reduction strategies. Despite the increasing use of AI in environmental assessment, economic feasibility remains a critical aspect in the evaluation of biogas systems. Among the few empirical contributions addressing this aspect, Vondra et al. (2019) propose a framework combining cost analysis and revenue modelling, to evaluate different operational scenarios and identifying the most influential variables. Overall, this cluster frames AI as a decision-

support tool at the strategic and system level, enabling the integration of environmental, economic, and operational dimensions.

3.2.2 Cluster 2- AI Applications for Biogas System Optimization

This research line is made up of 76 papers divided into 68 practical and 8 reviews papers. This cluster focuses on the use of AI to improve the operational performance of biogas systems. Cinar et al. (2021) highlight the significance of incorporating AI into biogas plants to facilitate sophisticated monitoring, prediction, and stabilization of processes. Likewise, Swami et al. (2023) propose an integrated framework for merging anaerobic digestion with data-driven technologies to optimize key operational parameters and enhance automation and real-time decision-making. Meanwhile, Cruz et al. (2022) examine the incorporation of ML algorithms into the modelling of anaerobic digestion processes and prove their efficiency in predicting biogas production. Practical research is mainly oriented towards the application of AI methods for estimating and optimizing biogas production. For instance, De Clercq et al. (2020) use algorithms such as the random forest and XGBoost models to analyze industrial-scale data and demonstrate their applicability and predictive accuracy. Salehi et al. (2022) create data-driven models relying on k-nearest neighbours and support vector machine techniques, which effectively predict biogas production and do not require specific kinetic information about the process. Moreover, Gonçalves Neto et al. (2021) implement artificial neural networks to estimate biogas production based on different organic substrates and evaluate the effect of temperature and other significant parameters. These studies confirm that AI-based models are effective tools for capturing the non-linear behavior of anaerobic digestion and improving process prediction and performance.

3.2.3 Cluster 3 - Advanced Computational Modelling and Optimization

Third stream contains 28 papers divided into 21 practical and 7 reviews papers. It focuses on the development and refinement of advanced algorithmic tools that enable predictive modelling and optimisation in biogas systems. Several review papers in this cluster focus on advanced computational and optimisation methods for modelling biogas systems. Ling et al. (2024) provide a comprehensive overview of ML algorithms, highlighting the benefits of hybrid models for improving prediction accuracy and process optimisation. Enitan et al. (2016) emphasise the role of evolutionary algorithms in solving complex, non-linear optimisation problems in anaerobic digestion systems. Wang et al. (2024) extend this perspective to the broader bioenergy sector, underlining the potential of algorithms for modelling, optimisation, and decision-making, while also addressing key challenges such as data limitations and model interpretability.

In addition to review studies, several practical contributions in this group of studies demonstrate the implementation of advanced computational techniques for modelling and optimisation of biogas systems. Gueguim Kana et al. (2012) combine artificial neural networks (ANN) with genetic algorithms (GA) to model and optimise co-substrate mixtures, achieving improved biogas yield and earlier production onset. Similarly, Abu Qdais et al. (2010) apply ANN and GA to real plant data, successfully predicting methane production and identifying optimal operating conditions. In addition, Kazemi et al. (2021) develop a data-driven framework based on soft sensors and ML models for fault detection, enhancing process monitoring and operational reliability.

4 Research Agenda/Future Research

Drawing upon the results from the content analysis, further research is outlined based on the three clusters (Table 4). Regarding the strategy-related and sustainable cluster, previous research focuses on utilizing AI for TEA and environmental assessment. While past literature predominantly applies an analytical view to the topic, there are no clear insights into how such information could be implemented in strategic decision-making. Hence, future research should concentrate on applying AI-based TEA in decision-making under uncertainty, focusing on investment decisions, scenario analysis, and strategic considerations. Furthermore, little is known about translating data-driven findings into decisions and their impact on decision quality and strategic behaviour in biogas facilities. In other words, further research should be undertaken to find out how the implementation of AI technologies changes decision-making processes to achieve more data and knowledge-driven decision-making. In the case of operational optimisation, most of the literature seems to focus on prediction and process efficiency. However, this view does not consider how AI can change the very nature of operational decision making in terms of its underlying logic. As such, future research will have to address the impact of AI integration into decision-making processes of anaerobic digestions and how it changes it. In particular, the issue needs to be explored as regards to whether AI allows for shifting from reactive operations into proactive ones and how this transformation is influenced by actual data availability and quality.

Lastly, regarding the advanced computational modelling stream, considerable progress has been achieved in terms of creating new modelling algorithms, but still the research area appears fragmented and rather performance oriented. Future research will have to move away from purely empirical and practical concerns towards integrative and theoretical ones and consider how advanced modelling techniques influence the way AD is represented and studied. Future research should explore whether hybrid and interlinked modelling frameworks allow for multi-level analysis of anaerobic digestion processes and how computational models help in generating knowledge and making decisions in digitalized

Table 4: Future research agenda on AI/ML application on biogas production.

Cluster	Open Research Questions
Cluster 1 – Strategic and Sustainability-Oriented AI Applications	· How can AI-driven TEA support investment decisions? · How are insights translated into decisions? · What is the role of AI in strategic planning?
Cluster 2 – Artificial Intelligence Applications for Biogas System Optimization	· How does the integration of AI reshape the decision-making logic of AD processes beyond traditional control systems? · To what extent does the adoption of AI shift biogas systems from reactive to predictive and proactive operational paradigms? · How do data availability and quality influence the effectiveness of AI-driven optimisation in real-world biogas plants?
Cluster 3 – Advanced Computational Modelling and Optimization	· How do advanced modelling approaches reshape the representation of AD processes, moving from mechanistic to data-driven system understanding? · To what extent can hybrid and integrated modelling frameworks enable a systemic and multi-level understanding of biogas systems? · How can computational models be designed not only for prediction, but also for supporting knowledge generation and decision-making in digitalised biogas systems?

5 Conclusions

This study introduces the notion that biogas can be considered as an energy company, in which digital technologies have significant influence on decision-making processes and performance achievements. From the perspective of managers, the results imply that implementation of AI in the production of biogas should not be restricted to enhancing the predictive precision or the efficiency of operations. Instead, AI should be used as a management tool to aid decision-making. In this sense, AI-powered techno-economic analysis will help make decisions and evaluate scenarios under conditions of uncertainty, while data-driven solutions will facilitate moving from a responsive to a more anticipatory and predictive mode of management. On the other hand, proper use of AI necessitates the formation of organisational competencies in data management and AI-human interaction. In terms of policymaking, the results show the necessity of enhancing institutional efforts to ensure digital transformation in the field of bioenergy. The government must take a critical role in supporting the use of innovative technology and developing AI tools, particularly in case of small-scale operations producing biogas. In addition, investments in modern digitalized technologies must be made to ensure effective implementation of the proposed technologies. It is also essential to mention that the digitalisation of biogas operations demands new skills, so training programs and knowledge transfer initiatives become crucial. Finally, the growing use of data driven tools brings up the issues associated with data protection and reliability of algorithms, so the development of corresponding regulatory policies is needed. Thus, there are several policy recommendations related to supporting digitalisation of biogas processes. First, it is necessary to design special incentives and offer funding programs for introducing innovations into practice. In addition, it is crucial to invest in modern digital technologies and develop regulations about data management and algorithm reliability.

Although this study makes certain contributions, it also has some limitations. Firstly, the literature analysis was based on Scopus and Web of Science. It is likely that other databases and sources contained papers which were also relevant for the topic and could have been included in the analysis. Secondly, even though the search string was carefully formulated, there might be other studies which could have been relevant but have not been captured due to the limited scope of the search strategy. In future studies, these limitations need to be addressed by analysing more databases and fine-tuning the search strategy. Additionally, more research is required to explore how the adoption of AI affects organisations, economies, and policies in practice and whether the conceptualization of biogas systems as digitalized energy enterprises can be validated empirically.

7 Ethics Declaration

The authors declare that no ethical approval was required for this research.

8 AI Declaration

In the preparation of this paper, AI tools, specifically ChatGPT and Gemini, were used solely to refine the English language and check grammar. These tools were not used to generate original content, ideas, or analysis. All intellectual work, arguments, and conclusions presented in this paper are the authors' own.

Main References

Ananda, A., Sujeeth, R.K. and Archana, S. (2025) Integrating automation in biomass transformation: opportunities, challenges, and future directions, *BioEnergy Research*, Vol. 18, No. 1, pp. 1–25.

Christou, M.L., Vasileiadis, S., Kalamaras, S.D., Karpouzas, D.G., Angelidaki, I. and Kotsopoulos, T.A. (2021) Ammonia-induced inhibition of manure-based continuous biomethanation process under different organic loading rates and associated microbial community dynamics, *Bioresource Technology*, Vol. 320, Article 124323.

De Clercq, D., Wen, Z., Fei, F., Caicedo, L., Yuan, K. and Shang, R. (2020) "Interpretable machine learning for predicting biomethane production in industrial-scale anaerobic co-digestion", *Science of the Total Environment*, Vol. 712, pp. 134574.

Kartal, M.T. (2022) The role of consumption of energy, fossil sources, nuclear energy, and renewable energy on environmental degradation in top-five carbon producing countries, *Renewable Energy*, Vol. 184, pp. 871–880.

Khan, M., Chuenchart, W., Surendra, K.C. and Khanal, S.K. (2023) Applications of artificial intelligence in anaerobic co-digestion: recent advances and prospects, *Bioresource Technology*, Vol. 370, Article 128501.

Overland, I., Juraev, J. and Vakulchuk, R. (2022) Are renewable energy sources more evenly distributed than fossil fuels?, *Renewable Energy*, Vol. 200, pp. 379–386.

Salehi, R., Yuan, Q. and Chaiprapat, S. (2022) "Development of Data-Driven Models to Predict Biogas Production from Spent Mushroom Compost", *Agriculture*, Vol. 12, No. 1090.

Salehian, P., Karimi, K., Zilouei, H. and Jeihanipour, A. (2013) Improvement of biogas production from pine wood by alkali pretreatment, *Fuel*, Vol. 106, pp. 484–489.

Swami, S., Suthar, S., Singh, R., Thakur, A.K., Gupta, L.R. and Sikarwar, V.S. (2025) Integration of anaerobic digestion with artificial intelligence to optimise biogas plant operation, *Environment, Development and Sustainability*, Vol. 27, No. 5, pp. 9773–9803.

Uddin, M.M. and Wright, M.M. (2023) Anaerobic digestion fundamentals, challenges, and technological advances, *Physical Sciences Reviews*, Vol. 8, pp. 2819–2837.

Ukpai, P.A. and Nnabuchi, M.N. (2012) Comparative study of biogas production from cow dung, cow pea and cassava peeling using 45 litres biogas digester, *Advances in Applied Science Research*, Vol. 3, No. 3, pp. 1864–1869.

Uranga-Valencia, L.P., Pérez-Álvarez, S., Gabriel-Parra, R., Chávez-Medina, J.A., Magallanes-Tapia, M.A., Sánchez-Chávez, E. and Castruita-Esparza, L.U. (2025) Biogas production from organic waste in the forestry and agricultural context: challenges and solutions for a sustainable future, *Energies*, Vol. 18, No. 12, Article 3174.

Vieira, A.S., Stewart, R.A., Lamberts, R. and Beal, C.D. (2020) Renewable energy and energy conservation area policy (REECAP) framework: a novel methodology for bottom-up and top-down principles integration, *Energy Strategy Reviews*, Vol. 32, Article 100544.

Waewsak, C., Nopharatana, A. and Chaiprasert, P. (2010) Neural-fuzzy control system application for monitoring process response and control of anaerobic hybrid reactor in wastewater treatment and biogas production, *Journal of Environmental Sciences*, Vol. 22, No. 12, pp. 1883–1890.

Wannapokin, A., Ramaraj, R. and Unpaprom, Y. (2017) An investigation of biogas production potential from fallen teak leaves (*Tectona grandis*), *Emerging Life Sciences Research*, Vol. 3, pp. 1–10.

Servant Leadership as a Relational Resource: Initial Insights from Qualitative Case Studies in Saudi Higher Education

Hayam Mohammed A Jabali[1], Ashish Malik[2], Philip J. Rosenberger III[1], Cathy Ying Xu[1]

[1] Newcastle Business School, University of Newcastle, Newcastle, Australia

[2] School of Business, UNSW Canberra, Canberra, Australia

Abstract

This study explores how servant leadership functions as a relational resource within resource-constrained higher education environments in Saudi Arabia. Drawing on the Job Demands–Resources (JD–R) framework, the research examines how leadership behaviours influence employees' experiences of high job demands, including workload pressures, administrative responsibilities, and institutional constraints. Using a qualitative multiple case study approach, data were collected through semi-structured interviews with leaders and followers across two cases – academic departments within a Saudi public university.

The study specifically focuses on how leadership is enacted and experienced in real-life organisational contexts, providing a nuanced understanding of how employees interpret leadership behaviours under conditions of sustained pressure and limited resources. This approach allows for the exploration of both leader and follower perspectives, offering a more comprehensive view of leadership dynamics within higher education settings.

The findings indicate that servant leadership does not reduce structural job demands but instead plays a critical role in buffering their negative effects by strengthening relational and social job resources such as trust, psychological safety and collaboration. Leadership behaviours including empathy, consultation and support were consistently identified as essential in helping employees cope with demanding work environments.

These behaviours were particularly significant in contexts where formal resources were limited, highlighting the role of leadership in compensating for structural constraints through relational support and interpersonal engagement. Employees reported that such leadership practices contributed to a more positive work environment, even in the presence of ongoing institutional challenges.

The study also highlights the importance of cultural context, demonstrating how leadership effectiveness is shaped by expectations related to hierarchy, respect and fairness.

These cultural dimensions influence not only how leadership behaviours are perceived but also how they are enacted. This requires leaders to balance authority with relational sensitivity to maintain credibility and trust within their teams.

Furthermore, social learning processes emerged as an important outcome, with employees relying on informal interactions and peer collaboration as key mechanisms for knowledge sharing. This finding underscores the importance of relational environments in facilitating learning and adaptation, particularly in settings where formal training opportunities may be limited. Leadership plays a central role in enabling these processes by fostering a culture of openness, collaboration and mutual support.

Overall, the findings contribute to the JD–R literature by positioning leadership as a central relational resource and provide contextually grounded insights into servant leadership within Saudi higher education.

Keywords: Servant leadership, Job Demands–Resources model, Saudi higher education, Qualitative case study, employee wellbeing.

Paper type: Academic Research Paper

1 Introduction

Leadership plays an important role in shaping how employees experience and respond to challenging work environments, particularly in higher education systems undergoing rapid transformation (Bakker & Demerouti, 2017). In Saudi Arabia, the higher education sector is facing increasing pressures associated with Vision 2030, including growing student numbers, accreditation requirements and heightened expectations for performance and quality (Al-shahrani, 2022). These pressures have significantly intensified job demands for both academic and administrative staff, often without a corresponding increase in resources (Akbar et al., 2023). Within this context, servant leadership has been proposed as an effective leadership approach due to its emphasis on support, empowerment and follower wellbeing (Eva et al., 2019). However, empirical research examining how servant leadership is enacted in practice within Saudi higher education, particularly through a Job Demands–Resources (JD–R) lens, remains limited (Gonaim, 2019). Preliminary insights from a localised, contextual and

cultural setting of Saudi Arabia highlight and offer a nuanced explanation of how servant leadership is enacted and how the key tenets of JD-R play out in this setting.

In addition to these structural pressures, recent developments in higher education have introduced increased expectations for research productivity, international collaboration and quality assurance processes, further intensifying the complexity of academic roles (Den Hartog & De Hoogh, 2024). These evolving expectations require employees to continuously adapt to changing institutional priorities, often under conditions of uncertainty and limited organisational support (Bakker et al., 2023). This highlights the growing importance of leadership approaches that can effectively support employees in navigating these challenges while maintaining wellbeing and performance (Eva et al., 2019).

Furthermore, examining servant leadership within the Saudi context provides an opportunity to better understand how leadership behaviours are shaped by cultural norms, organisational structures and institutional constraints. This is particularly important in high power-distance environments, where leadership expectations differ significantly from Western contexts (Shafai, 2020). Therefore, this study contributes to both theory and practice by offering contextually grounded insights into leadership within Saudi higher education (Den Hartog & De Hoogh, 2024).

This paper presents initial insights from an in-depth qualitative case study drawing on two cases (academic departments) to explore leadership, job demands, and job resources within resource-constrained higher education contexts. By examining servant leadership under high job demands, this study contributes to the IFKAD track on Leadership, Competencies and Culture for Intelligent Knowledge Management.

2 Methodology

The study adopts a qualitative case study approach grounded in an interpretive research paradigm (Bell et al., 2022). While the broader research project is based on a larger number of cases within Saudi higher education, this paper reports preliminary findings from two completed cases. Each case focuses on an academic department within a public university and draws on semi-structured interviews with four participants representing different organisational roles: faculty members, department chair and administrative staff. Across the two cases, the academic department served as the unit of analysis, enabling comparison of leadership practices and employee experiences under similar systemic pressures (Yin, 2018). Interview data were analysed using thematic analysis with the support of NVivo software to identify patterns related to leadership behaviours, job demands, job resources and employee experiences (Bell et al., 2022).

The use of qualitative case studies allows for a deeper exploration of how leadership is experienced within its real-life context, rather than assuming uniform effects across settings (Yin, 2018). This approach is particularly valuable in leadership research, where context plays a critical role in shaping behaviours and outcomes (Bell et al., 2022). Additionally, the inclusion of both leaders and followers ensures a more comprehensive understanding of leadership dynamics, reducing bias associated with single-perspective studies (Eva et al., 2019).

3 Findings

The findings indicate that both departments operate under extreme and sustained job demands. Participants consistently described high teaching loads, severe staff shortages, increasing administrative responsibilities and continuous accreditation pressures. Such demands align with previous research highlighting intensifying workloads and resource scarcity in higher education institutions (Mudrak et al., 2017). These pressures were further exacerbated by bureaucratic systems and institutional procedures that limited access to financial, human and developmental resources, increasing the risk of employee strain and burnout (Bakker et al., 2023).

Across both cases, participants repeatedly emphasised that these demands were not temporary or cyclical, but rather ongoing and embedded within the structure of the departments. One participant described the situation as "*constant pressure with no real break*," while another noted that "*the workload keeps increasing, but the number of staff does not change*." These reflections highlight the sustained nature of job demands and reinforce the perception that employees are operating in persistently resource-constrained environments.

A more detailed analysis shows that these demands are not isolated but interconnected, requiring employees to simultaneously manage teaching, administrative and research responsibilities under constrained conditions (Bakker et al., 2023). Faculty members described emotional exhaustion and reduced capacity for research engagement, with one participant noting that "*the most exhausting thing is the pressure combined with the lack of basic resources*", while administrative staff highlighted procedural rigidity and workload accumulation as key stressors (Akbar et al., 2023).

Importantly, the interconnected nature of these demands was experienced differently across the two cases. In one department, participants described a more acute sense of overload, where multiple responsibilities had

to be managed simultaneously with minimal support, leading to what one participant referred to as *"a feeling of always being behind."* In contrast, while the second department also experienced high demands, participants described a relatively more manageable environment, where some level of coordination and informal support helped mitigate the intensity of these pressures, as one participant noted that *"the level of cooperation is really high and everyone really does everything for the department"*. This suggests that while both cases share structural similarities, the experience of job demands varies in degree rather than in kind.

From a follower perspective, faculty members reported emotional and physical exhaustion, particularly when expected to maintain high standards of teaching and performance without adequate support or basic resources. Administrative staff similarly experienced pressure arising from rigid systems that prioritised compliance over employee wellbeing (Akbar et al., 2023). While the specific sources and intensity of pressure varied between the two departments, the persistence of high job demands emerged as a shared structural feature across both cases.

Participants further described how these pressures affected not only their performance but also their motivation and sense of professional fulfilment. One faculty member explained that *"you want to do good research, but you just don't have the time or energy,"* while another highlighted that *"most of the work becomes about meeting requirements, not improving quality."* These insights demonstrate how sustained job demands can shift employees' focus from meaningful engagement to task completion, reinforcing the strain associated with high-pressure environments.

Within this highly constrained and structured environment, servant leadership emerged as a critical relational resource rather than a mechanism for reducing objective workload. From the perspective of followers, participants described the department chair's leadership as respectful, consultative and supportive attributes they associated with servant leadership. Leadership behaviours, such as consulting staff before assigning tasks, offering emotional support and providing clarity around expectations, were perceived as essential in helping employees cope with overwhelming demands. These behaviours contributed to psychological safety, trust and commitment, as reflected in participants' experiences.

Participants provided concrete examples of how these leadership behaviours were enacted in practice. For instance, one participant noted that *"before assigning extra work, the chair always asks if we can manage it,"* while another highlighted that *"even when things are stressful, the chair explains clearly what is expected, which makes it easier to handle."* These examples illustrate how leadership behaviours are experienced directly by employees and how they influence day-to-day work experiences.

However, the effectiveness of these behaviours varied slightly across the two cases. In the more resource-constrained department, leadership was often described as a "buffer" that helped employees cope with stress, even when structural issues remained unresolved. In contrast, in the second department, leadership was perceived not only as supportive but also as enabling a more positive and collaborative work environment. This distinction suggests that servant leadership operates differently depending on the intensity of contextual constraints.

Across both cases, leadership behaviours were interpreted as meaningful forms of support, particularly when they addressed individual needs and acknowledged contextual constraints. Employees emphasised the importance of being heard, respected and involved in decision-making processes, highlighting the relational dimension of leadership as central to their work experience (Eva et al., 2019).

Participants frequently linked these relational practices to their sense of value and belonging within the department. As one participant expressed, *"when you feel that your opinion matters, it changes how you approach your work,"* while another noted that *"being included in decisions makes the workload feel more manageable."* These responses highlight the importance of relational dynamics in shaping not only how work is performed but also how it is experienced emotionally and psychologically.

Importantly, within this cultural context, these relational leadership practices serve as key leadership competencies through which leaders shaped a supportive and developmental culture that enabled informal knowledge sharing, peer problem-solving and context-specific learning practices. In this way, servant leadership contributed to the development of a local culture for effective knowledge management (Klein et al., 2013; Pellegrini et al., 2020).

The data further suggest that such practices were particularly important in contexts where formal systems and resources were limited. In both cases, employees relied heavily on peer support and informal communication to solve problems and share knowledge. One participant explained that *"we learn from each other more than from formal training,"* while another highlighted that *"most of the time, solutions come from discussions with colleagues."* These findings indicate that relational leadership not only supports individuals but also facilitates collective learning processes.

The findings also indicate that social learning processes were primarily informal, occurring through daily interactions, collaboration and shared experiences rather than structured training programs (Yang et al., 2023). This highlights the role of leadership in enabling environments where knowledge sharing and learning can emerge organically.

Across the two cases, these informal learning processes were consistently described as essential for navigating complex and changing work demands. However, the extent to which these processes were embedded varied. In one case, informal learning appeared as a necessity due to the absence of formal structures, whereas in the other, it was experienced as a complementary and valued aspect of the work environment. This again reflects variation in how similar dynamics are experienced across different contexts.

In sum, the findings demonstrate that while leadership does not eliminate structural constraints, it plays a critical role in shaping how these constraints are experienced by employees. By enhancing relational resources such as trust, support, and collaboration, leadership enables employees to better cope with high job demands and maintain engagement under pressure. Across both cases, leadership was consistently described as influencing the experience of workload rather than reducing the workload itself. Importantly, the two cases collectively illustrate that the impact of leadership is context-dependent, varying according to the intensity of job demands and the availability of resources, while still maintaining a consistent relational foundation across both settings.

4 Discussion

The findings extend the JD–R model by highlighting leadership as a central relational resource that shapes how employees experience job demands (Bakker et al., 2023). While traditional JD–R research focuses on job characteristics, this study demonstrates that leadership behaviours play a critical role in influencing employee wellbeing and engagement (Bakker & Demerouti, 2017).

More specifically, the findings suggest that leadership should not be viewed as an external factor within the JD–R framework, but rather as an embedded and dynamic resource that interacts directly with both job demands and job resources. This perspective contributes to extending the JD–R model by positioning leadership behaviours as an active mechanism that shapes how employees interpret, respond to and cope with workplace pressures (Bakker et al., 2023).

Servant leadership appears to function as a buffering mechanism that mitigates the negative effects of high job demands by enhancing relational resources such as trust, support and collaboration (Eva et al., 2019). Rather than directly reducing objective workload, these findings suggest that leadership influences the subjective experience of job demands, reinforcing the distinction between objective and perceived demands within the JD–R framework. This reinforces the importance of interpersonal dynamics in shaping employee experiences in demanding environments.

In this study, the buffering effect of servant leadership is particularly evident in situations where structural constraints cannot be altered. In such contexts, leadership behaviours do not reduce workload directly, but instead influence how workload is experienced by employees, making it more manageable and less psychologically taxing. This further supports the argument that relational resources play a critical role in sustaining employee functioning under pressure, extending existing JD–R assumptions about how resources operate in high-demand environments. This highlights the distinction between objective job demands and perceived job demands, with leadership playing a critical role in shaping the latter (Bakker & Demerouti, 2017).

Furthermore, the findings demonstrate that relational resources generated through servant leadership, such as psychological safety and trust, contribute to sustaining employee motivation and commitment even under prolonged pressure. This aligns with prior research, suggesting that social and relational factors are essential for maintaining engagement in high-demand environments (Eva et al., 2019).

The study also highlights the importance of cultural context in shaping leadership effectiveness. In Saudi higher education, leadership behaviours are interpreted through cultural expectations related to respect, hierarchy and fairness, which influence how employees evaluate leadership practices (Shafai, 2020).

Importantly, these cultural expectations do not merely influence perceptions of leadership but actively shape how leadership is enacted in practice. Leaders are required to navigate complex relational dynamics, including hierarchical sensitivity, age-related expectations and norms of respect, which influence their ability to exercise authority while maintaining positive relationships (Den Hartog & De Hoogh, 2024).

This suggests that servant leadership in this context is not a fixed or universal model, but rather an adaptive and contextually embedded practice that evolves in response to cultural and organisational constraints. As such, the effectiveness of servant leadership depends not only on the behaviours themselves but also on how well they align with the broader cultural and institutional environment(Shafai, 2020).

5 Practical Implications

The findings suggest that leadership development programs should focus on relational competencies, such as empathy, communication and support (Pellegrini et al., 2020). Additionally, institutions should create environments that encourage collaboration and informal learning processes (Yang et al., 2023).

From a practical perspective, these findings highlight the need for leadership training programs that move beyond technical and administrative skills to emphasise relational and interpersonal competencies. Developing leaders' ability to listen, communicate effectively and provide context-sensitive support is critical in environments characterised by high job demands and limited resources (Eva et al., 2019).

Moreover, higher education institutions should recognise the importance of informal learning and knowledge-sharing practices, which were found to play a central role in supporting employees. Creating spaces for collaboration, peer interaction and shared problem-solving can enhance organisational resilience and improve overall performance (Pellegrini et al., 2020).

At an organisational level, there is also a need to address structural barriers that limit leadership effectiveness. Policies and procedures that restrict flexibility, decision-making and resource allocation should be reviewed to enable leaders to provide more meaningful support to their teams (Akbar et al., 2023).

6 Limitations and Future Research

This study is based on two case studies, which may limit generalisability (Yin, 2018). Future research should expand the number of cases and could also look to incorporate quantitative approaches to validate the findings (Bell et al., 2022).

In addition to the limited number of cases, this study focuses on a specific institutional and national context, which may influence the transferability of the findings to other settings. While the insights provide valuable understanding of servant leadership in Saudi higher education, caution should be exercised when generalising these results to different cultural or organisational environments (Den Hartog & De Hoogh, 2024).

Future research could therefore use comparative studies across different cultural contexts to examine how servant leadership is interpreted and enacted in diverse environments. This would provide a more comprehensive understanding of the role of culture in shaping leadership effectiveness (Shafai, 2020).

Furthermore, integrating quantitative measures with qualitative insights could strengthen the validity of the findings and provide a more robust understanding of the relationship between leadership, job demands and employee outcomes (Bell et al., 2022).

7 Conclusion

Servant leadership acts as a relational resource within resource-constrained academic environments, supporting employees in managing high job demands. Rather than reducing workload, it enhances employees' capacity to cope through strengthening social and psychological resources (Bakker et al., 2023).

This study demonstrates that the value of servant leadership lies not in its ability to eliminate structural challenges, but in its capacity to reshape how these challenges are experienced by employees. By fostering trust, collaboration, and psychological safety, servant leadership enables employees to sustain engagement and performance despite ongoing pressures (Eva et al., 2019).

Overall, the findings contribute to a more nuanced understanding of leadership within the JD–R framework by emphasising the importance of relational dynamics, cultural context, and organisational constraints. This highlights the need for future research to further explore the intersection of leadership, culture, and employee wellbeing in complex organisational environments (Bakker et al., 2023).

8 Acknowledgements

The lead author would like to express sincere appreciation to her supervisors at the University of Newcastle for their continuous guidance, constructive feedback, and academic support throughout this research. The author also extends her gratitude to the co-supervisor from the University of Canberra (UNSW) for their valuable insights and contributions to the development of this study. In addition, the author is grateful to all participants who generously shared their time and experiences, which were essential to the completion of this research and significantly enhanced the quality and depth of the findings presented in this paper.

9 Ethics Declaration

This study adhered to ethical research standards and received approval by the university's Human Research Ethics Committee (approval number: H-2024-0357).

10 AI Declaration

AI tools were used in a limited and selective manner to support language refinement, editing and structuring of the manuscript. No AI tools were used in the conceptualisation, research design, data collection, data analysis or interpretation of the findings. Full responsibility is taken for the accuracy, validity and integrity of all content presented in this paper. All AI-assisted outputs were carefully reviewed, verified, and revised to ensure their appropriateness and alignment with the research objectives. The limitations of AI tools, including potential bias, inaccuracies and gaps in knowledge, are acknowledged, and it is confirmed that all final decisions and interpretations are entirely the author's own.

References

Akbar, H., Al-Dajani, H., Ayub, N., & Adeinat, I. (2023). Women's Leadership Gamut in Saudi Arabia's Higher Education Sector. *Gender, Work & Organization*, *30*(5), 1649-1675. https://doi.org/10.1111/gwao.13003

Al-shahrani, M. S. (2022). A review of higher education leadership in Saudi Arabia. *Pegem Journal of Education and Instruction*, *12*(4). https://doi.org/10.47750/pegegog.12.04.18

Bakker, A. B., & Demerouti, E. (2017). Job Demands-Resources Theory: Taking Stock and Looking Forward. *Journal of Occupational Health Psychology*, *22*(3), 273-285. https://doi.org/10.1037/ocp0000056

Bakker, A. B., Demerouti, E., & Sanz-Vergel, A. (2023). Job Demands–Resources Theory: Ten Years Later. *Annual Review of Organizational Psychology and Organizational Behavior*, *10*(1), 25-53. https://doi.org/10.1146/annurev-orgpsych-120920-053933

Bell, E., Bryman, A., & Harley, B. (2022). *Business Research Methods*. (6th ed.). Oxford University Press.

Den Hartog, D. N., & De Hoogh, A. H. B. (2024). Cross-Cultural Leadership: What We Know, What We Need to Know, and Where We Need to Go. *Annual Review of Organizational Psychology and Organizational Behavior*, *11*(1), 535-566. https://doi.org/10.1146/annurev-orgpsych-110721-033711

Eva, N., Robin, M., Sendjaya, S., van Dierendonck, D., & Liden, R. C. (2019). Servant Leadership: A Systematic Review and Call for Future Research. *The Leadership Quarterly*, *30*(1), 111-132. https://doi.org/10.1016/j.leaqua.2018.07.004

Gonaim, F. A. (2019). Leadership in Higher Education in Saudi Arabia: Benefits, Constraints and Challenges of Adopting Servant Leadership Model by Department Chairs. *International Journal of Education and Practice*, *7*(2), 101-111. https://doi.org/10.18488/journal.61.2019.72.101.111

Klein, A. S., Wallis, J., & Cooke, R. A. (2013). The impact of leadership styles on organizational culture and firm effectiveness: An empirical study. *Journal of Management & Organization*, *19*(3), 241-254. https://doi.org/10.1017/jmo.2013.34

Mudrak, J., Zabrodska, K., Kveton, P., Jelinek, M., Blatny, M., Solcova, I., & Machovcova, K. (2017). Occupational Well-being among University Faculty: A Job Demands-Resources Model. *Research in Higher Education*, *59*(3), 325-348. https://doi.org/10.1007/s11162-017-9467-x

Pellegrini, M. M., Ciampi, F., Marzi, G., & Orlando, B. (2020). The relationship between knowledge management and leadership: mapping the field and providing future research avenues. *Journal of Knowledge Management*, *24*(6), 1445-1492. https://doi.org/10.1108/jkm-01-2020-0034

Shafai, A. A. (2020). Servant Leadership in Higher Education in Saudi Arabia. In P. Mathew, J. Song, S. R. Ferch, L. C. Spears, Y. Alan, J. Arrais, G. Çelik, K. W. Chan, C. Crippen, & L. L. Hean (Eds.), *Global Servant-Leadership: Wisdom, Love, and Legitimate Power in the Age of Chaos* (pp. 253-271). Lexington Books/Fortress Academic.

Yang, J., Yu, T., Song, J., Cai, W., & Teng, R. (2023). A Job Demands-Resources Perspective on Servant Leadership and Employee Creativity. The Service Industries Journal, 1-23. https://doi.org/10.1080/02642069.2023.2238612

Yin, R. K. (2018). *Case Study Research and Applications: Design and Methods* (6th ed.). Sage Publication, Inc.

Retrieval-Augmented Generation in Business Applications: A Systematic Literature Review and Taxonomy

Beren Çay, Mario Miozza, Andrea De Mauro
Luiss Guido Carli University, Rome, Italy

Abstract

Retrieval-Augmented Generation (RAG) has become one of the most visible mechanisms for grounding large language model outputs in organizational knowledge sources. While the technical literature on RAG architectures is substantial, the evidence on how RAG is deployed inside organizations, and on the organizational consequences of those deployments, has not been systematically synthesized. This paper addresses that gap by treating RAG as a hybrid knowledge capability pairing a generative model with curated organizational knowledge and human oversight. The study reports a systematic literature review of peer-reviewed work on RAG in business and organizational contexts published between 2023 and 23 February 2026, covering Scopus and Web of Science in line with the PRISMA 2020 reporting framework. From 727 records after database-side filtering, 513 were screened and 41 retained. The studies were synthesized through the empirical-to-conceptual taxonomy development method of Nickerson, Varshney and Muntermann (2013) into a three-dimensional framework linking business application contexts, RAG architectures, and reported organizational consequences. Three findings stand out. First, adoption concentrates in knowledge-intensive support functions: enterprise knowledge retrieval, document intelligence, and decision support together account for around seventy percent of the corpus. Second, Classic RAG remains the most frequent single configuration, while Advanced and Agentic variants together appear in 18 studies and concentrate in deployments with more demanding task structures. Third, the most consistent finding is not a benefit but a limitation: residual output quality issues appear in 26 of 41 studies and recur across every architectural category and every application context, while data-quality dependence is the most frequent implementation challenge. The paper contributes a business-oriented taxonomy empirically distinct from architecture-first surveys, positioning RAG as a socio-technical capability whose organizational value depends on the surrounding data, governance, and oversight infrastructure. For practitioners, it supports deliberate decisions about where RAG delivers value, which architecture fits which task, and which conditions enable successful deployment.

Keywords: Retrieval-Augmented Generation, Artificial Intelligence, Knowledge Management, Business Applications, Systematic Literature Review

Paper type: Academic Research Paper

1 Introduction

Large language models (LLMs) have moved from a research topic to a routine component of organizational work, embedded in tasks from internal knowledge search to customer support and analytical decision assistance (Dwivedi et al., 2023). The operational difficulty is documented: LLMs generate fluent but unverifiable outputs — hallucination — and rely on a parametric memory fixed at training time, providing no clear provenance for its claims (Ji et al., 2023; Huang et al., 2025). In knowledge-intensive work where outputs feed decisions, are communicated externally, or are subject to audit, these are operational liabilities, not peripheral concerns.

Retrieval-Augmented Generation (RAG) was introduced by Lewis et al. (2020) for conditioning generation on retrieved external evidence rather than parametric memory alone. From a knowledge-management standpoint, the re-design matters less for benchmark performance than for the relationship between model and knowledge: the knowledge moves to an external store the organization can update, govern, inspect, and restrict (Gao et al., 2023; Fan et al., 2024). RAG sits inside the conceptual frame of Alavi and Leidner (2001) — knowledge management systems supporting creation, storage and retrieval, transfer, and application of organizational knowledge — most directly at retrieval and application stages.

The existing review literature is dominated by architecture-first technical surveys (Gao et al., 2023; Fan et al., 2024), where applications appear as illustrations rather than objects of analysis. Practitioner-oriented evidence on engineering challenges is more concrete (Barnett et al., 2024) but reported through software-engineering case studies rather than structured organizational synthesis. No business-oriented synthesis was identified that maps where RAG is used, which architectures are adopted, and what organizational benefits, limitations, and challenges are reported.

This paper provides that synthesis. Three research questions guide the review. RQ1: In which business functions and decision contexts is RAG applied? RQ2: What architectures are employed? RQ3: What benefits, limitations, and challenges are reported organizationally? The contribution is a three-dimensional taxonomy,

derived empirically from 41 peer-reviewed studies, linking application context, architectural configuration, and organizational consequence. It treats RAG as a hybrid-knowledge capability combining retrieval and generation with human oversight and curated organizational knowledge. Unlike architecture-first surveys (Gao et al., 2023; Fan et al., 2024), its primary axis is the organizational application context, not the technical pipeline; architectures and outcomes are read off that axis.

2 Theoretical Background

Three streams of literature are relevant. The first is the knowledge-management view that systems supporting organizational knowledge work must address the cycle of creation, storage and retrieval, transfer, and application (Alavi and Leidner, 2001). In knowledge-asset terms, RAG converts dispersed, weakly indexed organizational knowledge into a contextually retrievable, query-conditioned resource — without changing the underlying stock, but altering how it can be deployed. It functions at retrieval and application stages: not producing new organizational knowledge, but making existing knowledge accessible at the moment a question is asked. Dwivedi et al. (2023) argue that the impact of generative conversational AI depends on integration into work systems rather than capability alone.

The second stream concerns the technical anatomy of RAG. Parametric memory in a sequence-to-sequence model is paired with non-parametric memory accessed through a neural retriever; retrieved passages are incorporated into generation, producing more specific and factual outputs on knowledge-intensive tasks (Lewis et al., 2020). Gao et al. (2023) describe a progression from Naive RAG, through Advanced RAG (pre- and post-retrieval stages), to Modular RAG. Fan et al. (2024) document architectural diversification. Edge et al. (2024) introduce graph-based retrieval for sensemaking that requires reasoning over relational structure. The present paper organizes the architectural landscape into five categories recurring in the corpus: Classic RAG (Gao et al.'s Naive RAG), Advanced RAG, Agentic RAG, Hybrid RAG, Graph-Based RAG. The categories are structural alternatives whose suitability depends on task and underlying knowledge, not an ascending hierarchy.

The third stream is the still-thin organizational literature on RAG. Practitioner work has identified seven recurring failure points in engineered RAG systems and concluded that validation is meaningful only in operation, not in isolation (Barnett et al., 2024). Survey work on hallucination has shown that retrieval imprecision, irrelevant content, and failures to integrate retrieved evidence with generation can reintroduce the problems RAG was meant to address (Huang et al., 2025). What these contributions do not provide is a structured mapping of business contexts, architectures, and organizational consequences. That gap is the object of the present review.

3 Methodology

The study follows a systematic literature review methodology (Tranfield, Denyer and Smart, 2003), with reporting aligned to PRISMA 2020 (Page et al., 2021). The review was not pre-registered; the search strategy, eligibility criteria, screening, and extraction procedure are reported here to support replicability.

The literature search covered Scopus and Web of Science. The Boolean query — ("retrieval-augmented generation" OR "retrieval augmented generation" OR RAG) AND (business OR enterprise* OR organization*) — was applied to title, abstract and keyword fields in Scopus and to the Topic field in Web of Science. Business-side terms were kept broad to capture organizational applications across industries without biasing the corpus toward a single sector. The final search was conducted on 23 February 2026.

Scopus initially returned 912 records and Web of Science 469. After database-side filters for years (2023–2026), document types (articles, reviews, conference papers), and language (English), 494 Scopus and 233 Web of Science records were retained, giving 727 before deduplication. The 2023–2026 window was selected because organizational applications of RAG emerged after the broader diffusion of LLM-based systems. Records were exported to Rayyan, where 214 duplicates were removed, leaving 513 for screening.

Inclusion criteria, defined before screening, retained studies that (i) explicitly addressed RAG or closely related retrieval-augmented architectures, (ii) reported a concrete application, framework, prototype, case study, or empirical evaluation in an organizational or business context, and (iii) were peer-reviewed journal or conference papers. Exclusion criteria removed (i) exclusively technical studies with no organizational application, (ii) papers on generative AI or LLMs with a small or absent RAG component, (iii) non-academic publications, and (iv) secondary reviews without a substantive business-oriented case, framework, or application analysis.

Title-and-abstract screening reduced the corpus to 237 reports for full-text retrieval. Full texts were sought through the author's institutional subscriptions and Google Scholar; reports unavailable through either channel were classified as unretrieved, and the resulting accessibility bias is acknowledged in §5. Of these, 108 could not be retrieved. The remaining 129 were assessed in detail; 88 were excluded (3 not focused on RAG; 83 lacking

organizational contexts; 1 not a primary research study; 1 not peer-reviewed) and 41 retained. The full PRISMA 2020 flow is presented in Figure 1.

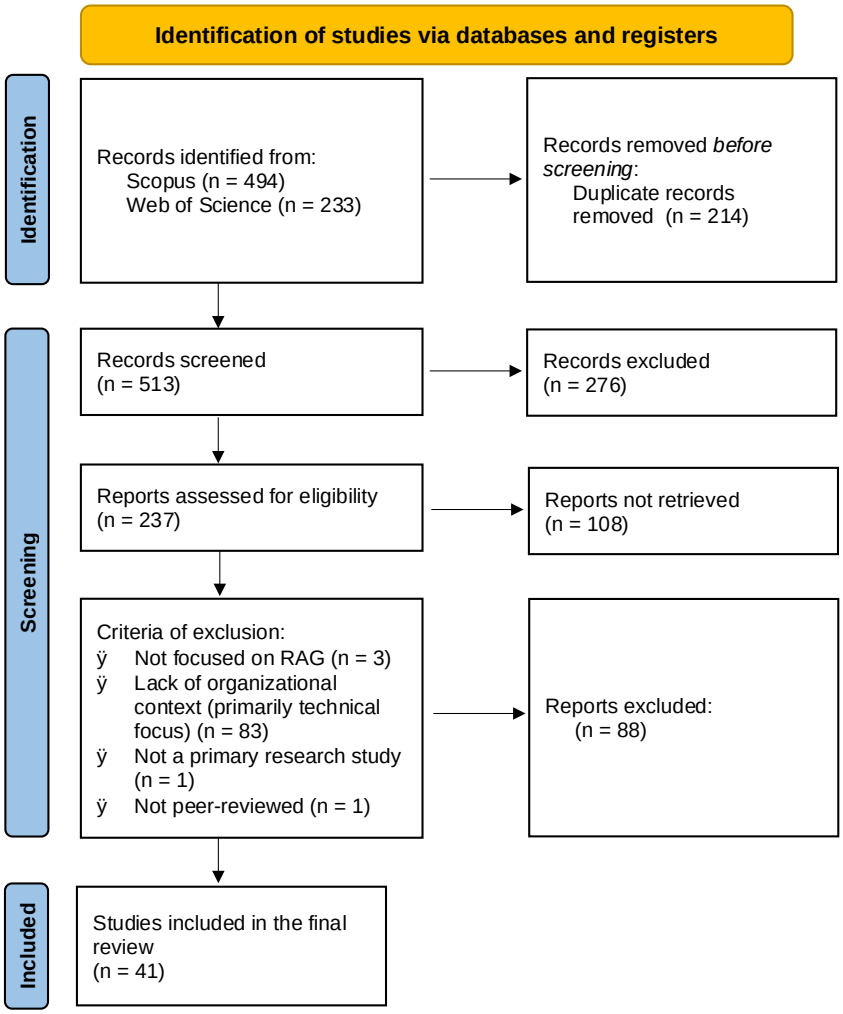


Figure 1: PRISMA 2020 flow diagram for the systematic literature search and screening process.

A formal quality appraisal was not performed: the corpus is heterogeneous in study type, and a single instrument would not apply evenly. Evaluation strength was recorded as a coded category at extraction (limitation *Weak or Missing Empirical Evaluation*, identified in 9 of the 41 studies and reported in §4.3) rather than aggregate scores. Coding and extraction were performed by a single reviewer; this and the absence of formal appraisal are acknowledged as methodological limitations. Each paper was reviewed using a purpose-built extraction instrument capturing sector, business application, RAG architecture, and reported organizational consequences. The taxonomy was developed following Nickerson, Varshney and Muntermann (2013), with the meta-characteristic defined as organisational use of RAG in business contexts.

4 Findings

The taxonomy is presented in Figure 2 and synthesized in Table 1.

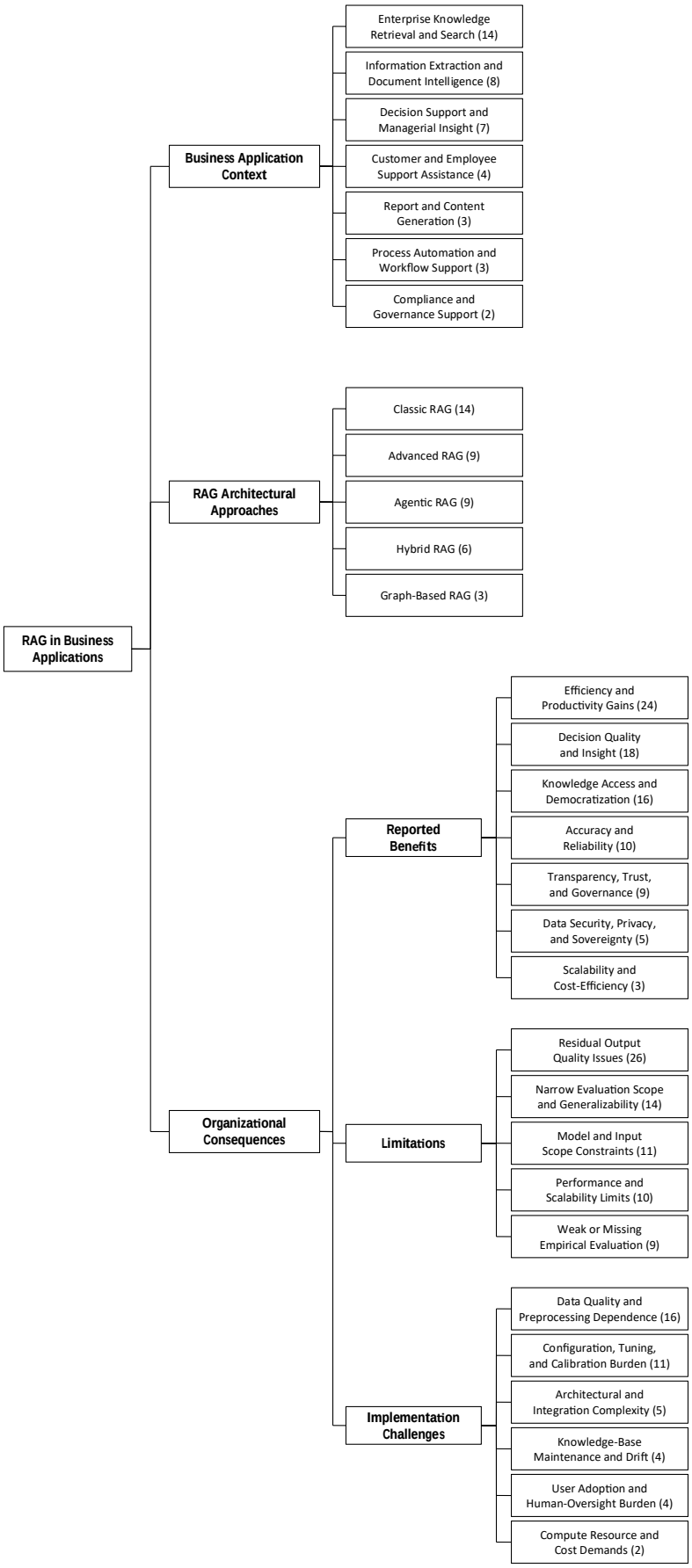


Figure 2: Three-dimensional taxonomy of Retrieval-Augmented Generation in business applications. Numbers in parentheses denote study counts within each category (n = 41 studies).

Table 1: Cross-dimensional synthesis of application context, observed architectures, and reported organizational consequences (n = 41 studies)

Business application context (n)	**Architectures observed (paper count)**	**Most frequently reported benefits**	**Most frequently reported limitations**	**Most frequently reported organizational challenges**	**Representative studies**
Enterprise Knowledge Retrieval and Search (14)	Hybrid (5); Advanced (3); Agentic (2); Classic (2); Graph-Based (2) — *all five architectures present*	Knowledge Access and Democratization (10); Efficiency and Productivity Gains (8); Accuracy and Reliability (4)	Residual Output Quality Issues (13); Model and Input Scope Constraints (4); Performance and Scalability Limits (3)	Data Quality and Preprocessing Dependence (7); Configuration, Tuning, and Calibration Burden (4)	Bordino et al. (2025); Choi et al. (2025); Eckroth et al. (2025); Tyndall et al. (2025); Viswanathan and Sasaki (2025)
Information Extraction and Document Intelligence (8)	Classic (4); Advanced (2); Agentic (2)	Efficiency and Productivity Gains (8); Decision Quality and Insight (7)	Narrow Evaluation Scope and Generalizability (5); Model and Input Scope Constraints (3); Performance and Scalability Limits (3); Weak or Missing Empirical Evaluation (3)	Data Quality and Preprocessing Dependence (4)	Arslan and Cruz (2024); Lin and Dang (2026); Vignesh et al. (2025)
Decision Support and Managerial Insight (7)	Classic (3); Advanced (2); Agentic (2)	Transparency, Trust, and Governance (4); Knowledge Access and Democratization (3); Decision Quality and Insight (3)	Narrow Evaluation Scope and Generalizability (4); Residual Output Quality Issues (3)	Data Quality and Preprocessing Dependence (4); Architectural and Integration Complexity (2)	Bi et al. (2025); Bernardi et al. (2024); Pitkäranta and Pitkäranta (2026); Colesanti Senni et al. (2025)
Customer and Employee Support Assistance (4)	Classic (3); Advanced (1)	Efficiency and Productivity Gains (2); Accuracy and Reliability (2); Knowledge Access and Democratization (2)	Residual Output Quality Issues (3); Weak or Missing Empirical Evaluation (2)	Configuration, Tuning, and Calibration Burden (2)	Della Sala (2025); Kruk, Herath and Choudhury (2025); Packowski et al. (2024); Orlowski, Knauff and Marquardt (2025)
Report and Content Generation (3)	Classic (1); Advanced (1); Hybrid (1)	Efficiency and Productivity Gains (2)	Narrow Evaluation Scope and Generalizability (2)	Configuration, Tuning, and Calibration Burden (2)	Alboaneen et al. (2025); Guan et al. (2025); Byun et al. (2025)
Process Automation and Workflow Support (3)	Agentic (2); Classic (1)	Efficiency and Productivity Gains (2)	Residual Output Quality Issues (3)	Configuration, Tuning, and Calibration Burden (1); User Adoption and Human-Oversight Burden (1)	Meddeb, Batti and Fathallah (2026) — Agentic; Rao et al. (2025) — Agentic; Cherednichenko and Maliarenko (2026) — Classic
Compliance and Governance Support (2)	Graph-Based (1); Agentic (1)	Decision Quality and Insight (2); Transparency, Trust, and Governance (2)	Each limitation reported once (n too low to identify dominant pattern)	Each challenge reported once (n too low to identify dominant pattern)	Sun, Luo and Li (2025) — Graph-Based; Pitkäranta and Pitkäranta (2024) — Agentic

Note. Architecture counts sum to row n and reflect dominant primary architecture coding. Benefits, limitations, and challenges are multi-coded per study; within-row counts may exceed n.

4.1 Dimension 1: Business application context

Across the 41 studies, seven contexts emerged. Enterprise Knowledge Retrieval and Search (14) is most frequent: banking knowledge bases (Bordino et al., 2025); automotive proprietary documentation (Choi et al., 2025); oil-and-gas community-of-practice assistants (Eckroth et al., 2025); enterprise resource planning (Viswanathan and Sasaki, 2025); privacy-preserving public-sector configurations (Tyndall et al., 2025). Information Extraction and Document Intelligence (8) synthesizes unstructured documents — manufacturing specifications (Lin and Dang, 2026); supply-chain mapping from financial filings (Jackson et al., 2026). Decision Support and Managerial Insight (7) foregrounds analytical orientation — climate-disclosure assessment (Colesanti Senni et al., 2025); human-aligned AI governance in banking (Pitkäranta and Pitkäranta, 2026). Together these account for around seventy percent of the corpus, indicating concentration in knowledge-intensive support functions rather than autonomous high-stakes settings. The remaining four — Customer and Employee Support (4), Report and Content Generation (3), Process Automation and Workflow Support (3), Compliance and Governance Support (2) — are present but limited.

4.2 Dimension 2: RAG architectural approaches

Five architectural categories appeared. Classic RAG, where a query triggers retrieval and retrieved content is added to the prompt (Lewis et al., 2020), is most frequent (14 of 41). Advanced RAG and Agentic RAG share second place (9 each). Advanced RAG extends the pipeline through re-ranking, validation, and self-reflective refinement, for tasks demanding greater precision or verifiability — bill-of-materials digitization (Lin and Dang, 2026); supply-chain mapping (Jackson et al., 2026). Agentic RAG embeds RAG in agent-based loops for planning, tool use, multi-step execution (Meddeb, Batti and Fathallah, 2026; Pitkäranta and Pitkäranta, 2024, 2026). Hybrid RAG (6) combines dense semantic with lexical or structured retrieval for heterogeneous enterprise knowledge bases. Graph-Based RAG (3) uses an explicit knowledge graph (Edge et al., 2024) for relational domains: regulatory compliance (Sun, Luo and Li, 2025); physical-internet logistics (Naganawa, Hirata and Yamada, 2025); ERP ontology (Viswanathan and Sasaki, 2025). The categories are not hierarchical: the prevalence of Classic RAG suggests substantial value (knowledge access, response grounding, reduced hallucination risk) is achievable through direct configurations. Architectural-selection rationale is rarely explicit; the distribution reflects outcomes.

4.3 Dimension 3: Organizational consequences

Nine of 41 studies did not explicitly report implementation challenges (coded as *Not explicitly reported*); categorical counts use the remaining 32, multi-coded. The most frequently reported benefit is Efficiency and Productivity Gains (24 studies): a 12.5× speedup in manufacturing document processing (Lin and Dang, 2026); 85% reduction in supply-chain rule-modification time (Meddeb, Batti and Fathallah, 2026); 20–30% latency improvement in enterprise retrieval (Khaliq and Adebayo, 2026). Decision Quality and Insight (18) and Knowledge Access and Democratization (16) follow, the latter where deployments reduce dependence on senior expertise (Eckroth et al., 2025). Accuracy and Reliability (10) and Transparency, Trust, and Governance (9) are secondary but consequential, the governance cluster appearing in regulated and high-accountability settings (Sun, Luo and Li, 2025; Pitkäranta and Pitkäranta, 2024; Alboaneen et al., 2025; Colesanti Senni et al., 2025). Data Security, Privacy, and Sovereignty (5) and Scalability and Cost-Efficiency (3) are least frequent.

The most frequently reported limitation is Residual Output Quality Issues, in 26 of 41 studies. This captures incomplete, vague, inconsistent, or insufficiently grounded outputs persisting even when generation is augmented by retrieval — hallucinations surviving RAG (Cherednichenko and Maliarenko, 2026); long-context handling difficulties (Choi et al., 2025); retrieval imprecision (Bordino et al., 2025). Its near-universal appearance across all five architectural categories and seven application contexts is the single most consistent finding. Narrow Evaluation Scope and Generalizability (14), Model and Input Scope Constraints (11), Performance and Scalability Limits (10), and Weak or Missing Empirical Evaluation (9) round out the limitations.

The most frequently reported implementation challenge is Data Quality and Preprocessing Dependence (16), capturing the degree to which RAG performance depends on the cleanliness, completeness, and structure of the underlying knowledge base. Configuration, Tuning, and Calibration Burden (11) is the effort to select embedding models, set retrieval thresholds, refine prompts, tune re-ranking, and adapt per domain. Architectural and Integration Complexity (5), Knowledge-Base Maintenance and Drift (4), User Adoption and Human-Oversight Burden (4), and Compute Resource and Cost Demands (2) complete the dimension.

4.4 Cross-dimensional observations

Two patterns are organizationally consequential. First, Residual Output Quality Issues appear across every architectural category and every application context — sophistication does not, in this evidence, eliminate them. This is hard to reconcile with the assumption that elaborate configurations (agentic, graph-based, hybrid) automatically resolve quality failures. Second, Data Quality and Preprocessing Dependence appears across all architectural categories, more frequent in elaborate ones (5 of 9 Advanced, 4 of 9 Agentic, against 3 of 14 Classic). Bin sizes are small and the pattern suggestive, but consistent with deployment success sitting outside the architecture, in surrounding organizational and data infrastructure. A third pattern: Graph-Based RAG appears in domains with explicit relational structure rather than a distinctive application category, illustrating architectural choice following knowledge-domain structure rather than a general upgrade path.

5 Discussion

Read against the research questions, the taxonomy supports reading RAG as a hybrid knowledge capability, not a self-contained technical artefact. For RQ1, adoption concentrates in the knowledge-management problem space — inaccessibility of dispersed institutional knowledge, processing of unstructured text, grounding judgements in traceable sources — consistent with the observation that re-use of dispersed knowledge is a main bottleneck in knowledge work (Alavi and Leidner, 2001). For RQ2, the prominence of Classic RAG alongside meaningful shares for Advanced and Agentic configurations in more demanding tasks suggests that, at this stage of adoption, the simplest fit-for-purpose configuration tends to win. For RQ3, the persistence of residual quality issues, together with data-quality dependence, supports a socio-technical reading: RAG's value depends on surrounding infrastructure rather than residing in the model.

For organizations in regulated and sustainability-relevant settings, the corpus offers a specific signal. Transparency, Trust, and Governance concentrates in compliance, governance, and disclosure-assessment applications (Sun, Luo and Li, 2025; Colesanti Senni et al., 2025; Alboaneen et al., 2025). The case for RAG is context-dependent: in operations-intensive settings the argument is efficiency; in regulated and ESG-relevant settings, grounded and auditable reasoning. Framing RAG only as hallucination reduction fails to capture where its value concentrates in practice.

For practitioners, three implications follow. First, data readiness is the primary adoption prerequisite, and its importance persists — perhaps intensifies — as architectural sophistication grows: a well-maintained knowledge base enables more reliable outcomes from a modest configuration than an elaborate architecture on poorly prepared documentation. Second, architecture should follow task structure: an Agentic system for a use case adequately served by Classic RAG adds configuration and governance burden without proportionate benefit. Third, human review and governance are operating requirements, not safeguards.

The study has methodological limitations. Coding was performed by a single reviewer; the absence of inter-rater reliability testing means classification reflects interpretive judgement. The review was not pre-registered. The search was restricted to two databases (Scopus and Web of Science), to peer-reviewed journal and conference papers, and to publications in English. 108 reports could not be retrieved and may have introduced accessibility bias. No formal quality appraisal was performed. The 2023–2026 window captures the active period of organizational RAG adoption but is short.

6 Conclusion

This paper reviews the RAG literature in business and organizational contexts and develops a three-dimensional taxonomy linking application context, architectural approach, and organizational consequence. Adoption concentrates in knowledge-intensive support functions; Classic RAG dominates while Advanced and Agentic configurations cluster in demanding deployments; benefits coexist with a persistent residual-quality limitation regardless of architecture, and with data-infrastructure dependence at least as strong in elaborate as in simple architectures. RAG is positioned as a hybrid knowledge capability whose value depends on surrounding data, governance, and oversight infrastructure. The next research step is comparative, longitudinal, governance-focused evidence from real deployments. For practice, the taxonomy supports decisions about where RAG delivers value, which architecture fits which task, and which conditions enable deployment.

7 Ethics declaration

This study did not require ethical clearance. No human or animal subjects were involved, and no personal or sensitive data were collected, processed, or analyzed.

8 AI declaration

The author used AI-assisted tools during the preparation of this paper. Generative AI was used to support language editing and structural refinement in selected sections. All conceptual work, methodological decisions, screening, coding, taxonomy development, interpretation of findings, and final wording were carried out by the author, who reviewed and verified all AI-assisted text and takes full responsibility for the content.

References

Alavi, M. and Leidner, D.E. (2001) "Review: Knowledge Management and Knowledge Management Systems: Conceptual Foundations and Research Issues", *MIS Quarterly*, Vol. 25, No. 1, pp. 107–136.

Alboaneen, D., Alhajri, S., Alhajri, K. et al. (2025) "Siyasat: AI-Powered AI Governance Tool to Generate and Improve AI Policies According to Saudi AI Ethics Principles", *Computers*, Vol. 14, No. 11, 452.

Arslan, M. and Cruz, C. (2024) "Business-RAG: Information Extraction for Business Insights", in *Proceedings of ICSBT 2024*, SciTePress, Dijon, pp. 88–94.

Arslan, M., Munawar, S. and Cruz, C. (2024a) "Business insights using RAG–LLMs: a review and case study", *Journal of Decision Systems*, advance online publication.

Arslan, M., Munawar, S. and Cruz, C. (2024b) "Exploring Business Events using Multi-source RAG", *Procedia Computer Science*, Vol. 246, pp. 4534–4540.

Arslan, M., Munawar, S. and Cruz, C. (2024c) "Sustainable Digitalization of Business with Multi-Agent RAG and LLM", *Procedia Computer Science*, Vol. 246, pp. 4722–4731.

Arslan, M., Munawar, S. and Cruz, C. (2026) "Business-RAG: Advancing Enterprise Information Extraction with Llama and DeepSeek", in *Smart Business Technologies, ICSBT 2024, CCIS Vol. 2589*, Springer, Cham, pp. 17–35.

Arslan, M., Munawar, S. and Riaz, Z. (2025) "Empowering SMEs with SustainWater Bot to advance urban water sustainability", *Sustainable Cities and Society*, Vol. 132, 106793.

Avogadri, S., Alzetta, G. and Russo, D. (2026) "Discovery Omnia: Dynamic RAG for Enhanced Patent Analysis and Systematic Innovation", in *World Conference of AI-Powered Innovation and TRIZ Methodology, TFC 2025, IFIP AICT Vol. 774*, Springer, Cham, pp. 87–99.

Barnett, S., Kurniawan, S., Thudumu, S. et al. (2024) "Seven Failure Points When Engineering a Retrieval Augmented Generation System", in *Proceedings of the IEEE/ACM 3rd International Conference on AI Engineering – Software Engineering for AI*, ACM, New York, pp. 194–199.

Bernardi, M.L., Casciani, A., Cimitile, M. and Marrella, A. (2024) "Conversing with business process-aware large language models: the BPLLM framework", *Journal of Intelligent Information Systems*, Vol. 62, pp. 1607–1629.

Bi, F., Cao, D., Wang, Z. et al. (2025) "RBDQ: A Reliable LLM-based Text-to-SQL System for Business Data Queries", in *Companion Proceedings of the ACM Web Conference 2025*, ACM, New York, pp. 95–103.

Bordino, I., Di Iorio, F., Galliani, A. et al. (2025) "UniAsk: AI-powered search for banking knowledge bases", in *Proceedings of the 28th International Conference on Extending Database Technology*, OpenProceedings, pp. 1057–1065.

Byun, G., Lee, S., Choi, N. and Choi, J.D. (2025) "Secure Multifaceted-RAG: Hybrid Knowledge Retrieval with Security Filtering", *Information*, Vol. 16, No. 9, 804.

Caglar, E., Keles, M., Kutanoglu, M. and Demir, M. (2025) "ELIF: An End-to-End Architecture for an Observable and Continuously Learning AI Assistant", *The Eurasia Proceedings of Science, Technology, Engineering & Mathematics (EPSTEM)*, Vol. 35, pp. 46–54.

Cherednichenko, O. and Maliarenko, V. (2026) "AI-Based Efficient Automation of Decision Logic Representation", in *Smart Business Technologies, ICSBT 2025, CCIS Vol. 2666*, Springer, Cham, pp. 303–315.

Choi, N., Byun, G., Chung, A. et al. (2025) "Reference-Aligned Retrieval-Augmented Question Answering over Heterogeneous Proprietary Documents", in *Proceedings of CIKM 2025*, ACM, New York, pp. 5626–5633.

Chu, S. and Liu, J. (2025) "Based on BERT-GPT-GNN converged architecture: intelligent generation engine for complex SQL queries in business intelligence", *Discover Artificial Intelligence*, Vol. 5, No. 1, 147.

Colesanti Senni, C., Schimanski, T., Bingler, J., Ni, J. and Leippold, M. (2025) "Using AI to assess corporate climate transition disclosures", *Environmental Research Communications*, Vol. 7, No. 2, 021010.

Della Sala, J. (2025) "Technical Customer Service Support with RAG Fine Tuned LLaMA 3", *The International FLAIRS Conference Proceedings*, Vol. 38, No. 1.

Dwivedi, Y.K., Kshetri, N., Hughes, L. et al. (2023) "Opinion Paper: 'So what if ChatGPT wrote it?' Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy", *International Journal of Information Management*, Vol. 71, 102642.

Eckroth, J., Freitag, D., Keefe, J. et al. (2025) "Building a Virtual Member of a Community of Practice", in *Proceedings of CIKM 2025*, ACM, New York, pp. 5650–5657.

Edge, D., Trinh, H., Cheng, N. et al. (2024) "From Local to Global: A Graph RAG Approach to Query-Focused Summarization", *arXiv preprint* arXiv:2404.16130.

Fan, W., Ding, Y., Ning, L. et al. (2024) "A Survey on RAG Meeting LLMs: Towards Retrieval-Augmented Large Language Models", in *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, ACM, New York, pp. 6491–6501.

Gao, Y., Xiong, Y., Gao, X. et al. (2023) "Retrieval-Augmented Generation for Large Language Models: A Survey", *arXiv preprint* arXiv:2312.10997.

Guan, X., Ye, Z., Cao, B. and Fan, J. (2025) "Summary report auto-generation based on hierarchical corpus using large language model", *Displays*, Vol. 89, 103055.

Hong, Y. and Kim, D. (2025) "Innovative Applications of RAG-Enhanced Small LLM for Closed-Domain Q&A", *International Journal of Innovative Computing, Information and Control*, Vol. 21, No. 2, pp. 481–490.

Huang, L., Yu, W., Ma, W. et al. (2025) "A Survey on Hallucination in Large Language Models: Principles, Taxonomy, Challenges, and Open Questions", *ACM Transactions on Information Systems*, Vol. 43, No. 2, Article 42.

Jackson, I., Sáenz, M.J., Ivanov, D. and Ma, B.J. (2026) "Supply chain mapping through retrieval-augmented generation: applications to the electronics industry", *Journal of the Operational Research Society*, advance online publication.

Jeong, C. (2023) "A Study on the Implementation of Generative AI Services Using an Enterprise Data-Based LLM Application Architecture", *Advances in Artificial Intelligence and Machine Learning*, Vol. 3, No. 4, pp. 1588–1618.

Ji, Z., Lee, N., Frieske, R. et al. (2023) "Survey of Hallucination in Natural Language Generation", *ACM Computing Surveys*, Vol. 55, No. 12, Article 248.

Keshireddy, S.R. (2025) "Retrieval-Augmented Generation Techniques in Oracle APEX Improving Contextual Responses in AI Assistants", *Archives for Technical Sciences*, Vol. 33, No. 2, pp. 253–270.

Khaliq, A. and Adebayo, K.J. (2026) "DKC-LLM: Dynamic Knowledge Caching for Large Language Models in Business Applications", *IEEE Access*, Vol. 14, pp. 22318–22334.

Kruk, F., Herath, S. and Choudhury, P. (2025) "BanglAssist: A Bengali-English Generative AI Chatbot for Code-Switching and Dialect-Handling in Customer Service", in *Extended Abstracts of CHI 2025*, ACM, New York, Article 93.

Lewis, P., Perez, E., Piktus, A. et al. (2020) "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks", in *Advances in Neural Information Processing Systems*, Vol. 33, pp. 9459–9474.

Lin, Y.-C. and Dang, J.-F. (2026) "From product specifications to smart manufacturing: Generative AI-driven process innovation for BOM digitization and its applications", *Advanced Engineering Informatics*, Vol. 71, 104451.

Meddeb, S., Batti, S. and Fathallah, H. (2026) "BPMN DMN decision table generation based on agentic AI for critical applications", *Business Process Management Journal*, advance online publication.

Naganawa, H., Hirata, E. and Yamada, A. (2025) "Implementing a Knowledge Management System with GraphRAG: A Physical Internet Example", *Electronics*, Vol. 14, No. 24, 4948.

Nagpal, R., Usua, U., Palacios, R. and Gupta, A. (2025) "FairRAG: A Privacy-Preserving Framework for Fair Financial Decision-Making", *Applied Sciences*, Vol. 15, No. 15, 8282.

Nickerson, R.C., Varshney, U. and Muntermann, J. (2013) "A method for taxonomy development and its application in information systems", *European Journal of Information Systems*, Vol. 22, No. 3, pp. 336–359.

Orlowski, M., Knauff, E. and Marquardt, F. (2025) "Towards Personal Assistants for Energy Processes Based on Locally Deployed LLMs", in *Proceedings of ICAART 2025, Volume 3*, SciTePress, pp. 695–706.

Packowski, S., Halilovic, I., Schlotfeldt, J. and Smith, T. (2024) "Optimizing and Evaluating Enterprise Retrieval-Augmented Generation (RAG): A Content Design Perspective", in *Proceedings of ICAAI 2024*, ACM, New York, pp. 162–167.

Page, M.J., McKenzie, J.E., Bossuyt, P.M. et al. (2021) "The PRISMA 2020 statement: an updated guideline for reporting systematic reviews", *BMJ*, Vol. 372, n71.

Pitkäranta, T. and Pitkäranta, L. (2024) "Bridging Human and AI Decision-Making with LLMs: The RAGADA Approach", in *Proceedings of ICEIS 2024 – Volume 1*, SciTePress, pp. 812–819.

Pitkäranta, T. and Pitkäranta, L. (2026) "Using Agentic LLM Architecture to Align Human and AI Decisions", in *Enterprise Information Systems, ICEIS 2024 Revised Selected Papers, LNBIP Vol. 566*, Springer, Cham, pp. 170–194.

Rao, K., Coviello, G., Mellone, G., De Vita, C.G. and Chakradhar, S. (2025) "XPF: Agentic AI System for Business Workflow Automation", in *Proceedings of HPDC 2025*, ACM, New York.

Seabra, A., Cavalcante, C., Nepomuceno, J. et al. (2026) "Enhanced Document and Database Integration for Advanced Question-Answering in Enterprise Contract Management with LLMs and Agents", in *Knowledge Discovery, Knowledge Engineering and Knowledge Management, IC3K 2024, CCIS Vol. 2703*, Springer, Cham, pp. 150–173.

Su, Y., Mirafzal, R. and Stal-Le Cardinal, J. (2025) "Exploring LLM-based agents for need analysis of knowledge management practice", *Proceedings of the Design Society*, Vol. 5 (ICED25), pp. 1675–1684.

Sun, J., Luo, Z. and Li, Y. (2025) "A Compliance Checking Framework Based on Retrieval Augmented Generation", in *Proceedings of COLING 2025*, Association for Computational Linguistics, pp. 2603–2615.

Tranfield, D., Denyer, D. and Smart, P. (2003) "Towards a Methodology for Developing Evidence-Informed Management Knowledge by Means of Systematic Review", *British Journal of Management*, Vol. 14, No. 3, pp. 207–222.

Tyndall, E., Wagner, T., Gayheart, C., Some, A. and Langhals, B. (2025) "Feasibility Evaluation of Secure Offline Large Language Models with Retrieval-Augmented Generation for CPU-Only Inference", *Information*, Vol. 16, No. 9, 744.

Vignesh, G., Harikrishnan, P.M., Reddy, S. et al. (2025) "IDPFlow: A No-Code Agentic Framework for Multimodal Intelligent Document Processing", in *Proceedings of ACM Multimedia 2025*, ACM, New York, pp. 14359–14360.

Viswanathan, A. and Sasaki, F. (2025) "Combining Knowledge Graphs and Retrieval Augmented Generation for Enterprise Resource Planning", in *Advances in Information Retrieval, ECIR 2025, LNCS Vol. 15576*, Springer, Cham, pp. 111–115.

AI-Enabled Coordination Framework to address Polycrisis: Human–Machine Collaboration and Regenerative Capacity

Alessandro Margherita[1], Tina Comes[2], Gianluca Elia[1], Gianluca Solazzo[1], Istvan Hoffman[3], Jay Liebowitz[4]

[1] University of Salento, Lecce, Italy;
[2] Delft University, Delft, Netherlands & Institute for Terrestrial Infrastructure Protection, DLR, Germany;
[3] Eötvös Loránd University, Budapest, Hungary;
[4] Thomas Jefferson University, Philadelphia, Pennsylvania, USA

Abstract

Contemporary crises, ranging from pandemics and climate-related disasters to geopolitical conflicts, are increasingly interconnected and transboundary, giving rise to *polycrises* characterised by cascading effects, high uncertainty, and fragmented information. These dynamics challenge traditional crisis management and call for coordination mechanisms capable of aligning heterogeneous stakeholders under time pressure and asymmetric knowledge. Although Artificial Intelligence (AI), particularly generative AI, is gaining prominence in crisis management, it is still predominantly framed as a tool for prediction or automation. This paper argues that such a view overlooks AI's broader potential as a socio-technical coordination mechanism that augments human capabilities by supporting shared sensemaking and knowledge integration across organisational and institutional boundaries. Adopting a design science research approach combined with conceptual development, this paper proposes a reference framework that conceptualises AI as an embedded coordination infrastructure. The framework integrates three interrelated components. First, it specifies five AI-enabled coordination roles (Sentinel, Broker, Verifier, Orchestrator, and Scribe) aligned with core knowledge management processes (sensing, integration, validation, application, retention). These roles are derived by linking coordination dependencies (flow, fit, share) to knowledge processes that become critical during polycrises and are conceived as overlapping mechanisms embedded in human-AI interaction and organisational practices. Second, it proposes a four-layer architecture that operationalises these roles across heterogeneous environments, spanning data/provenance, knowledge processing, coordination enablement, and governance and learning. Third, the study introduces the construct of AI-enhanced regenerative coordination capacity, articulated through three pillars of crisis readiness: dependency visibility, coordination mechanism agility, and collective learning and institutionalisation. To make the concepts concrete without claiming empirical validation, the framework is examined through structured vignettes derived from recurring crisis patterns and assessed via an analytical, vignette-based evaluation. A common specification template supports traceability from roles to layers to outputs and enables assessment of coverage, governability (provenance, contestability, bounded authority), and explanatory usefulness. Practically, the framework informs the design of interoperable, transparent, and adaptive crisis coordination systems that strengthen informed trust, organisational learning, and cross-boundary collaboration in polycrisis settings.

Keywords: Coordination; Crisis Management; Human-AI Collaboration; Knowledge Management; Sensemaking

Paper type: Academic Research Paper

1 Introduction and Problem Statement

Contemporary crises, such as pandemics, climate-related disasters and geopolitical conflicts, are increasingly characterised by their interconnected and transboundary nature. These events unfold across organisational, institutional, and territorial boundaries, creating a complex web of interdependencies. As a result, disruptions propagate across systems, generating cascading effects that amplify their impact and duration. Such dynamics challenge the effectiveness of traditional crisis management approaches, which are typically designed for more linear, predictable, and bounded scenarios (e.g., Comes et al., 2020; Reichstein et al., 2025).

These evolving conditions are often described as "polycrises," emphasising the interaction of multiple, overlapping disruptions. In such contexts, the challenge extends beyond responding to a single event to coordinating multiple actors operating under different institutional logics, priorities, and information regimes. Decision-makers must navigate radical uncertainty, where the nature of the problem may be unclear, while dealing with fragmented and rapidly evolving information. Time pressure further intensifies these constraints, requiring timely yet informed decisions despite incomplete knowledge. Consequently, effective crisis response increasingly depends on coordination mechanisms capable of aligning distributed actors, supporting shared understanding, and enabling continuous adaptation as situations evolve (Margherita et al., 2024).

In recent years, Artificial Intelligence (AI), particularly generative AI, has emerged as a promising resource in crisis management due to its ability to process large volumes of heterogeneous data and support decision-

making in real time (Otal et al., 2024; Pantiris et al., 2025). However, existing research largely frames AI in terms of automation and predictive accuracy. While valuable, this perspective underestimates its potential in complex, multi-actor environments, where the primary challenge lies not only in prediction but in coordination (Domfeh & Dancy, 2025).

Specifically, this prevailing view overlooks AI's capacity to function as a collaborative coordination mechanism that augments, instead of replacing, human capabilities. In crisis settings, where knowledge is distributed and continuously evolving, fostering shared understanding becomes critical. AI systems can support this process by integrating diverse data sources, translating information across domains, and making interdependencies more visible, thereby enabling more effective knowledge sharing and alignment among stakeholders (Li et al., 2025; Rai et al., 2025).

Building on this perspective, this paper reconceptualises AI as a socio-technical coordination mechanism embedded within organisational and inter-organisational processes. While AI is often treated as a passive tool, it is conceptualised as an active component that structures relational complexity, shapes how actors interpret information, and supports collective sensemaking during crises. Drawing on coordination theory, the study emphasises the management of interdependencies among activities and actors, as well as the role of shared knowledge in enabling coordinated action (Nespeca et al., 2020; Okhuysen & Bechky, 2009).

By positioning AI within this broader socio-technical framework, the paper addresses the following research question:

How can AI facilitate coordination in highly complex and uncertain environments?

In this view, the research develops a design-oriented perspective that explains how AI can enhance the capacity of organisations and networks to collectively respond to, learn from, and adapt to crises.

2 Background

The analysis developed in this study focuses on crises in which neither human intelligence nor artificial intelligence alone is sufficient to ensure effective decision-making and coordination. Instead, these contexts call for human–AI collaboration, where complementary strengths are combined to achieve superior outcomes (Li et al., 2025; Rai et al., 2025). Human actors contribute contextual understanding, ethical judgment, and the ability to interpret ambiguity, while AI systems provide scalability and the capacity to process heterogeneous data streams in near real time (Pantiris et al., 2025; Otal et al., 2024). The value of these complementarities is amplified in polycrisis settings, where knowledge is distributed in multiple sources and across organisations, and must be continuously integrated and updated (Nespeca et al., 2020).

Such settings are typically characterised by interrelated conditions that intensify coordination complexity. First, severe time pressure constrains deliberation and increases the need for rapid sensemaking and timely action (Margherita et al., 2024). Second, information is fragmented and asymmetrically distributed across stakeholders operating under different institutional logics, technological infrastructures, and governance regimes, often producing inconsistent situational interpretations (Comes et al., 2020; Nespeca et al., 2020). Third, task, resource, and information interdependencies evolve dynamically, making it difficult to anticipate how local decisions propagate across the system (Okhuysen & Bechky, 2009). Finally, crisis decisions are high-stakes and consequential, raising the premium on reliable, adaptive, and cross-boundary coordination (Reichstein et al., 2025).

A core implication of these conditions is that crisis coordination requires not only pattern detection but also interpretable reasoning about event chains and cascading effects. Purely correlation-driven analytics may identify regularities yet remain fragile when the crisis departs from historical patterns or when spurious relationships drive misleading signals (Pearl & Mackenzie, 2018). In this respect, causal reasoning approaches can strengthen crisis decision-making by supporting counterfactual thinking and scenario exploration, helping stakeholders distinguish plausible mechanisms from coincidental associations and better anticipate the consequences of alternative interventions (Pearl & Mackenzie, 2018; Li et al., 2025). However, the practical value of such reasoning depends on how it is embedded into collaborative processes that support shared interpretation and coordinated action across organisational boundaries (Rai et al., 2025; Margherita et al., 2024).

Within this broader perspective, generative AI (and large language models in particular) is increasingly relevant for crisis coordination because it can function both as an enabling infrastructure and as an active coordination mechanism. As infrastructure, it supports cross-domain translation and synthesis of heterogeneous inputs, and it can generate coordination artefacts such as situation reports, task assignments, and protocol recommendations. As a coordination mechanism, it can contribute to monitoring evolving conditions, highlighting emerging interdependencies, routing information to relevant actors, and supporting the production of shared representations needed for collective sensemaking. These capabilities align with recent surveys on

LLMs for coordination and cooperation tasks (Zhu et al., 2024), and complement work in crisis management information systems and actor-centred coordination (Nespeca et al., 2020; Benaben et al., 2020).

At the same time, applying generative AI in crisis contexts raises constraints that are not peripheral but architectural. LLMs can hallucinate and generate plausible yet incorrect outputs (Sandvik, 2025), a risk that is magnified under time pressure and cognitive bias (Comes, 2016). They can also be brittle under novelty, when crises deviate from patterns represented in training data (Comes, 2026), and their performance is strongly dependent on the quality, timeliness, and governance of underlying data. These limitations motivate human-centred design requirements, such as human-in-the-loop and override mechanisms, provenance and auditability of recommendations, explicit uncertainty communication, and governance arrangements that preserve accountability and legitimacy in high-stakes decisions (Shneiderman, 2020; SAPEA, 2025; Comes, 2024).

Overall, this background supports the need to move beyond viewing AI as an isolated decision-support tool and instead conceptualise it as part of a socio-technical system in which coordination quality depends on how human and artificial capabilities are integrated within workflows, knowledge processes, and governance structures (Hevner et al., 2004; Gregor & Hevner, 2013). Existing research on AI in crisis management still predominantly frames AI in terms of automation, prediction, or isolated decision support, under-specifying its role in managing interdependencies and shaping cross-boundary sensemaking and knowledge integration (Harika et al., 2024; Karinshak, 2024; Piller, 2024). This study addresses this gap by reconceptualising generative AI as an embedded coordination infrastructure and by developing a design-oriented framework that links coordination roles, architectural layers, and regenerative readiness in polycrisis contexts.

3 Framework of AI-Enabled Coordination in Crisis Management

3.1 Research approach

This study adopts a design-oriented approach that combines Design Science Research (DSR) (Hevner et al., 2004; Gregor & Hevner, 2013) with conceptual development (Jaakkola, 2020). This positioning reflects the exploratory nature of the work, which aims to advance theoretical understanding while proposing a structured way to think about AI-enabled coordination in crisis contexts, instead of delivering a fully validated prescriptive solution. The primary outcome of the study is a conceptual artefact in the form of a reference framework for AI-enabled coordination. The framework integrates three interrelated components: (i) a set of AI-enabled coordination roles aligned with key knowledge management processes (sensing, integration, validation, application, and retention), (ii) a layered architecture describing how these roles can be enacted across heterogeneous actors and systems, and (iii) a set of design-oriented considerations that clarify how such configurations may support coordination under conditions of uncertainty and interdependence.

The development of the artefact is informed by an iterative process of conceptual synthesis. Specifically, insights from coordination theory, knowledge management, human–AI collaboration, and crisis management literature are combined to identify and relate core constructs, such as interdependencies, distributed sensemaking, and AI-supported knowledge processes. As opposed to following a fully formalized build–evaluate cycle, the approach emphasises the gradual refinement and internal coherence of the framework through theory integration. Evaluation is limited and primarily analytical in nature. The framework is examined in terms of its internal consistency and its ability to capture salient characteristics of polycrisis contexts, including fragmented information, dynamic interdependencies, and high uncertainty. Scenario construction and evaluation are detailed in the next section.

3.1.1 Scenario construction and analytical evaluation

The operationalisation of the framework consists of structured scenario vignettes represented in Figure 1 and based on recurring crisis patterns (e.g., pandemics, climate-related disruptions, infrastructural failures), which are not used as empirical validation but as heuristic illustrations. Each scenario follows a common specification template: (1) *context and threat dynamics* (including escalation and cascading effects); (2) *stakeholder constellation and institutional logics* (public agencies, operators, NGOs, communities); (3) *salient flow–fit–share dependencies* (task, resource, and information interdependencies); (4) *information conditions* (fragmentation, latency, and asymmetry); (5) *expected coordination breakdown risks* (e.g., overload, misalignment, coordination-information bubbles); and (6) *candidate role–layer interventions and coordination artifacts* (e.g., dependency maps, shared situation reports, escalation triggers).

Scenarios are selected to ensure cross-boundary coordination demands across at least two domains (e.g., health and infrastructure), explicit cascading effects, high-stakes decision-making under time pressure, and informational asymmetry.

Evaluation is analytical and scenario-based: it is assessed whether the framework (a) covers the core polycrisis conditions, (b) supports traceability from roles to layers to outputs, (c) renders dependencies explicit and actionable, (d) remains governable through provenance, contestability, and bounded authority, and (e) provides explanatory usefulness for identifying where coordination value is generated and where breakdowns are likely to occur.

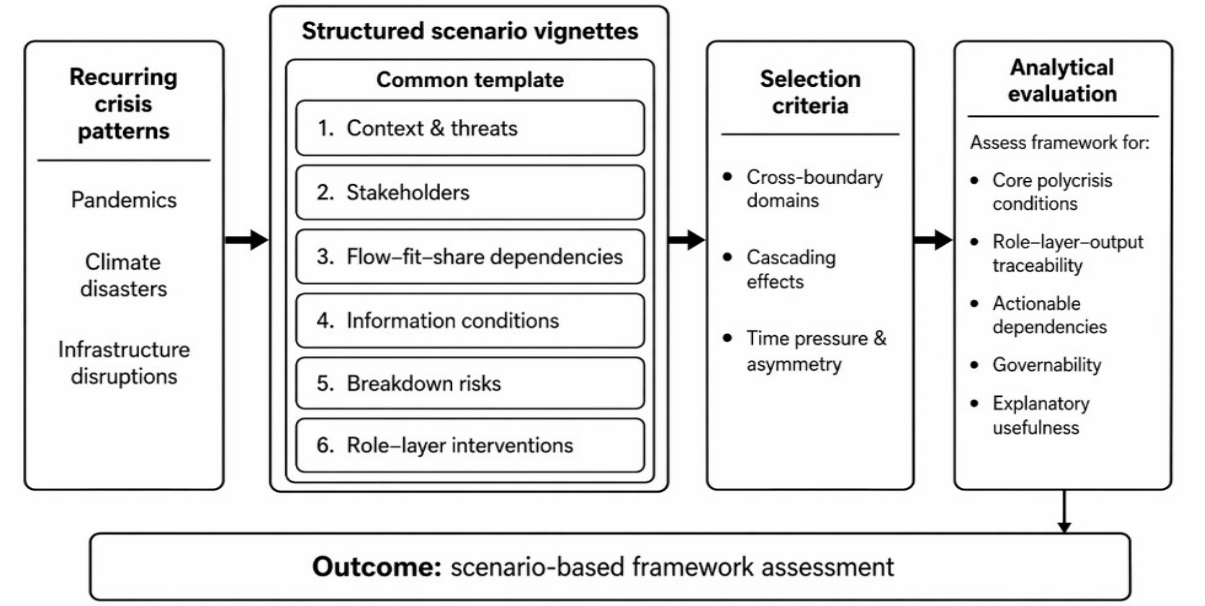


Figure 1. Scenario construction and evaluation process

3.2 AI-enabled coordination roles

This study specifies five AI-enabled coordination roles as the functional core of the proposed framework. The roles are derived by linking coordination dependencies (flow, fit, and share) to five knowledge management processes that become critical under polycrisis conditions—sensing, integration, validation, application, and retention. In this view, AI contributes to coordination by supporting the acquisition, transformation, and use of distributed knowledge across heterogeneous actors, thereby shaping how interdependencies are perceived and managed during crises. Given the dynamic nature of roles and responsibilities in crises (Turoff et al., 2004), the roles are not conceived as discrete or sequential stages; rather, they operate as overlapping and mutually reinforcing mechanisms embedded in human–AI interaction and organisational practices. Table 1 summarises the five roles, their coordination function, and illustrative applications in crisis management. By structuring AI-enabled coordination around knowledge processes, the framework provides a parsimonious way to link technical capabilities with organisational needs. The table offers a synthetic representation of these roles, their coordination functions, and illustrative applications in crisis and emergency management settings.

Table 1. Synopsis of AI-enabled Coordination Roles, related KM processes and Examples.

AI-Enabled Coordination Role	Knowledge Management Process	Description (Coordination Function)	Illustrative Example in Crisis/Emergency Management
Sentinel (Signal Detection & Monitoring)	Sensing	Identify and aggregate signals from heterogeneous data sources to support shared situational awareness, in line with distributed sensemaking and information processing (e.g. Comes et al., 2020).	*Real-time aggregation of satellite data, social media posts, and emergency calls to detect emerging flood zones or wildfire spread.*
Broker (Knowledge Integration & Translation)	Integration	Generate synthetic data and combine or translate distributed information across actors and domains, enabling shared representations and coordination across boundaries (e.g. Okhuysen & Bechky, 2009; Nespeca et al., 2020).	*Filling data gaps and integration of hospital capacity data, mobility patterns, and public health reports during a pandemic to create a unified operational picture.*

AI-Enabled Coordination Role	Knowledge Management Process	Description (Coordination Function)	Illustrative Example in Crisis/Emergency Management
Verifier (Consistency Checking & Causal Assessment)	Validation	Assess reliability, reconcile inconsistencies, and support causal reasoning to distinguish robust signals from spurious correlations (e.g. Rai et al., 2025).	*Cross-validation of field reports and sensor data to confirm infrastructure damage, or use of causal models to assess whether a supply shortage is driven by logistics disruption or demand spikes.*
Orchestrator (Decision Support & Coordination Enablement)	Application	Support decision-making and alignment of interdependent actions under uncertainty, consistent with coordination as the management of dependencies (e.g. Margherita et al., 2024).	*Scenario-based allocation of emergency resources (e.g., ambulances, shelters) and coordination of evacuation plans across agencies.*
Scribe (Learning & Knowledge Institutionalisation)	Retention	Capture and codify knowledge to enable organisational learning and the adaptation of coordination practices over time (e.g. Pimentel & Veliz, 2025).	*Post-crisis analysis of response actions to update emergency protocols, improve preparedness plans, and train future responders.*

3.3 Layered digital architecture for coordination

While the coordination roles defined in Section 3.2 clarify what AI-enabled systems do to support crisis coordination, an operating architecture is needed to explain how these functions are enacted across heterogeneous socio-technical environments. Building on a conceptual synthesis of research on digital infrastructures, multi-level coordination, and AI-enabled ecosystems (Nespeca et al., 2020; Cui et al., 2024), we propose four interdependent layers: (a) Data & Signal, (b) Knowledge Processing, (c) Coordination Enablement, and (d) Governance & Learning. The design logic is that coordination in polycrisis settings requires the progressive transformation of raw signals into validated and actionable knowledge, coupled with feedback and accountability mechanisms that sustain trust and adaptation over time.

The *Data & Signal Layer* provides the foundation of the architecture by acquiring and preparing heterogeneous inputs (e.g., sensor networks, satellite feeds, social media streams, institutional databases, and field reports) for downstream processing. At this level, AI-enabled functions primarily support the *Sentinel* role through ingestion, normalisation, time-stamping, aggregation, and initial relevance filtering. Core outputs include curated signal streams, basic quality indicators, and provenance metadata required for subsequent integration and validation. A central design constraint concerns data reliability under crisis conditions as well as the identification and bridge of data gaps, especially in polycrisis contexts, where sources may be delayed, incomplete, contradictory, or strategically manipulated, which makes uncertainty tagging and provenance preservation necessary at the point of capture.

The *Knowledge Processing Layer* transforms aggregated signals into integrated and contextually meaningful representations that can be shared across heterogeneous actors. It operationalises the *Broker and Verifier* roles by supporting cross-domain translation, semantic/ontological alignment, and consistency checking, including causal assessment when appropriate. The primary output is a validated, multi-perspectival situational knowledge base, accompanied by uncertainty and provenance cues, which enables shared interpretation and provides a reliable input to the coordination enablement functions in the upper layer.

The *Coordination Enablement Layer* translates validated knowledge into coordinated action. It operationalises the *Orchestrator* role by generating coordination artefacts such as situational summaries, resource allocation options, protocol recommendations, dependency alerts, and escalation triggers. This layer also manages dependency tracking: it maintains a continuously updated model of task, resource, and information interdependencies among actors, which allows it to identify coordination gaps, flag bottlenecks, and trigger adaptive responses when conditions change. A key architectural feature at this level is the presence of explicit human-in-the-loop mechanisms. Given the high-stakes nature of crisis decisions and the known limitations of AI systems with respect to hallucination and brittleness under novel conditions (Sandvik, 2025), the layer is designed to present options and recommendations instead of autonomous decisions.

The *Governance & Learning Layer* operates as the meta-level of the architecture. It operationalises the *Scribe* role in the Retention Process by performing three distinct functions. First, it monitors the quality and coherence of outputs generated by the lower layers, flagging inconsistencies or deviations from established protocols. Second, it captures and codifies knowledge generated during crisis episodes, structuring it into reusable artefacts such as updated response templates, revised causal models, and annotated case repositories. Third, it

manages access, accountability, and transparency across the system, including audit of AI-generated recommendations and compliance with applicable governance frameworks.

The four layers are linked by two interface mechanisms. Upward flows carry progressively abstracted knowledge, from raw signals to actionable coordination outputs. Downward flows carry feedback from higher layers to recalibrate lower-layer operations (e.g. when governance monitoring identifies a systematic bias in the signal detection function or when learning outputs update the ontologies used in knowledge integration). Figure 2 provides a synoptic representation of the architecture.

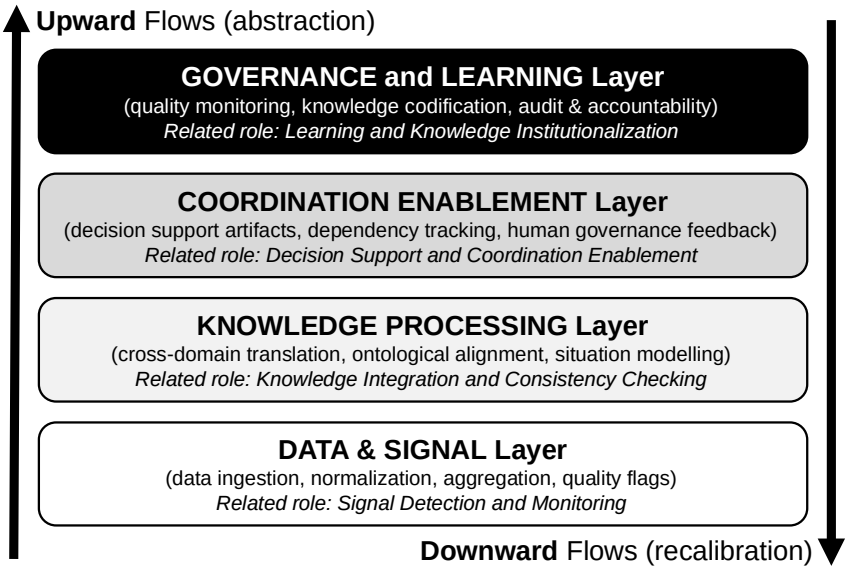


Figure 2. Structure of Layered Architecture

3.4 AI-enhanced coordination capacity

AI actively shapes collective sensemaking processes by influencing what information becomes visible, how it is interpreted, and how it circulates across actors (Comes, 2026). We introduce the concept of AI-enhanced regenerative coordination capacity as a conceptual extension of a higher-level theoretical abstraction of the impact of coordination roles and architecture on organisational readiness.

This capacity is defined along three interrelated pillars: a) dependency visibility, b) coordination mechanism agility, and c) collective learning and institutionalisation. *Dependency visibility* refers to the ability to make interdependencies observable and interpretable across distributed actors. In polycrisis contexts, the failure to perceive dependencies (between organisations, resources, information systems, and decision processes) is a primary source of coordination breakdowns. AI systems, particularly those operating at the Data & Signal and Knowledge Processing layers of the architecture, enhance this visibility by continuously mapping and updating dependency structures in near-real time, making latent interdependencies explicit and presenting them in formats that support human interpretation and action.

Coordination mechanism agility captures the capacity to dynamically reconfigure coordination structures in response to evolving conditions. Static coordination plans are inherently brittle in polycrisis environments where the nature and boundaries of the problem shift rapidly. AI-enabled systems contribute to agility by monitoring changes in dependency structures, identifying misalignments between existing coordination arrangements and emerging operational realities, and supporting the rapid adaptation of protocols, resource allocations, and communication flows. This pillar is most directly operationalised through the Coordination Enablement Layer and the Decision Support role.

Collective learning and institutionalisation emphasise the role of AI in supporting knowledge accumulation and adaptation over time. Crisis response does not end with the resolution of an immediate event; its value also lies in the organisational learning it generates. The Governance & Learning Layer operationalises this pillar by structuring the capture, validation, and codification of knowledge produced during crisis episodes into reusable organisational resources that enhance future preparedness.

Together, these three dimensions frame crisis readiness as a dynamic and regenerative capability, as opposed to a static condition, one that is continuously rebuilt through the interaction of AI-enabled functions,

human judgment, and institutional processes. Figure 3 provides a synoptic representation of the three pillars and their relationships.

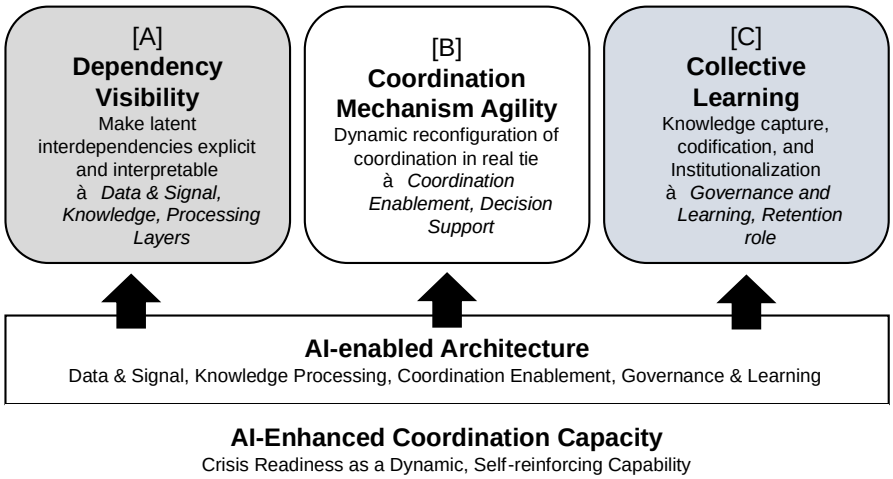


Figure 3. Coordination Capacity Pillars

4 Discussion and Implications

4.1 Academic implications

This study contributes to coordination and crisis management research by specifying how AI-enabled systems can be theorised as coordination mechanisms in polycrisis settings. Moving beyond considering AI as an external decision-support tool, the framework clarifies where and how coordination value is produced through the coupling of functional roles (Sentinel, Broker, Verifier, Orchestrator, Scribe) and architectural layers. This specification supports a testable and design-relevant account of human–AI collaboration in crisis coordination, with particular relevance for knowledge-intensive, cross-boundary settings.

A first contribution lies in extending research on AI-enabled systems and digital infrastructures. Existing approaches often emphasise performance improvements within bounded organisational settings, whereas this study conceptualises AI as an infrastructural layer that operates across organisational and institutional boundaries. By proposing a reference framework tailored to polycrisis contexts, the paper highlights how AI can structure interactions among heterogeneous stakeholders, enabling alignment despite fragmented information and divergent institutional logics. This perspective moves beyond system-centric views and contributes to an emerging stream of research on ecosystem-level coordination and AI-enabled socio-technical architectures.

Second, the paper advances coordination science by introducing a mechanism-based taxonomy of AI-enabled coordination roles. Unlike traditional coordination models, which treat technology as an external support, this framework positions AI as an active participant in the management of interdependencies. By linking coordination roles to knowledge management processes and to flow–fit–share dependencies, the study provides a more granular and operationalizable account of how coordination unfolds in hybrid human–AI environments. This represents a meaningful extension of coordination theory, particularly in contexts characterised by uncertainty, distributed knowledge, and dynamic interdependencies.

Third, the study contributes by integrating functional, structural, and dynamic perspectives into a unified framework. The combination of coordination roles (functional dimension), layered architecture (structural dimension), and regenerative coordination capacity (dynamic dimension) offers a multi-level analytical lens that is largely absent in existing models. This integration enables a more comprehensive understanding of how coordination mechanisms are designed, how they operate in practice, and how they evolve.

Finally, the paper links roles and architecture to AI-enhanced regenerative coordination capacity, articulated through dependency visibility, coordination agility, and collective learning and institutionalisation. The scenario-based analytical evaluation in Section 3.1.1 supports this linkage by tracing role–layer mechanisms to coordination outputs and readiness pillars, enabling comparative designs and evaluation beyond technical performance.

4.2 Practitioner implications

For practitioners in crisis management, the framework suggests that effective adoption of AI requires designing coordination infrastructures, not isolated analytics tools. A practical priority is interoperability by design: data and signal ingestion must be paired with shared semantic standards and dependency representations that can be interpreted across organisations. This implies investing not only in technical integration but also in shared vocabularies, role definitions, and accountability arrangements that allow information to travel without losing meaning or legitimacy.

A second implication concerns decision support under bounded authority. In high-stakes environments, AI should be configured to produce structured options, trade-offs, and escalation triggers (rather than autonomous decisions) while making assumptions and uncertainty explicit. Human-in-the-loop mechanisms should therefore be treated as architectural features, coupled with provenance and audit trails that support post hoc accountability and learning.

A third implication concerns organisational learning as a design target. Governance and learning mechanisms should capture crisis episodes as reusable knowledge assets (e.g., updated playbooks, annotated cases, revised ontologies) linked to decision provenance, enabling institutionalisation through training and standard updates. The scenario template in Section 3.1.1 can be used to stress-test designs ex ante by specifying stakeholders, dependencies, information conditions, and breakdown risks.

5 Conclusions

This paper highlights that the value of Artificial Intelligence (AI) in crisis and polycrisis contexts lies fundamentally in its ability to support regenerative coordination outcomes. These outcomes include enhanced adaptability, continuous learning, and the progressive strengthening of collective crisis response capabilities over time. By enabling the integration of distributed knowledge, the alignment of heterogeneous actors, and the dynamic reconfiguration of coordination mechanisms, we presented an integrated framework of AI-enabled coordination as an evolving process and a foundational element in the development of resilient socio-technical systems capable of evolving through disruption and uncertainty. This perspective shifts the focus of crisis management from reactive response toward the cultivation of long-term adaptive capacity embedded within coordination infrastructures.

The study acknowledges some limitations. The proposed framework has been analytically developed and validated through conceptual reasoning and scenario-based demonstration, but it has not been empirically tested. Moreover, the illustrative scenarios, while useful for demonstrating internal coherence and applicability, do not capture the full complexity and variability of real-world crises. As such, the findings should be interpreted as design-oriented and exploratory and not as conclusive.

Future research should therefore focus on empirical validation of the proposed framework through case studies, simulations, or experimental approaches, as well as on comparative analyses across different crisis domains and governance contexts. In addition, further investigation is needed to better understand the potential risks and unintended consequences associated with AI-enabled coordination, including issues related to bias, opacity, over-reliance on automated systems, and the governance of human–AI interaction in high-stakes environments. Addressing these challenges will be critical to ensuring that AI-enabled coordination infrastructures remain not only effective but also trustworthy, accountable, and aligned with broader societal objectives.

6 Ethics Declaration

No ethics issues or requirements for this research.

7 AI Declaration

No AI was used to realize this paper.

References

Benaben, F., Fertier, A., Montarnal, A., Mu, W., Jiang, Z., Truptil, S., ... & Lamothe, J. (2020). 'An AI framework and a metamodel for collaborative situations: Application to crisis management contexts'. Journal of Contingencies and Crisis Management, Vol 28, No. 3, 291-306.

Comes, T. (2016, March). Cognitive biases in humanitarian sensemaking and decision-making: lessons from field research. In 2016 IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA), 56-62, IEEE.

Comes, T. (2024). AI for crisis decisions. Ethics and Information Technology, 26(1), 12.
Comes, T. (2026). 'Sensemaking AI: Introducing a research and design agenda for human–AI networks'. EPJ Data Science.
Comes, T., Van de Walle, B. and Van Wassenhove, L. (2020) 'The coordination-information bubble in humanitarian response: theoretical foundations and empirical investigations', Production and Operations Management, Vol. 29, No. 11, pp. 2484–2507.
Cui, Q., Shen, L. and Sun, Y. (2024) 'How ecosystems coordinate architectures and AI in humanitarian operations? A configurational view', Technological Forecasting and Social Change, 201, 123456.
Domfeh, E.A. and Dancy, C.L. (2025) 'Human-AI use patterns for decision-making in disaster scenarios: a systematic review', in 2025 IEEE International Symposium on Technology and Society (ISTAS), pp. 1–10.
Gregor, S. and Hevner, A.R. (2013) 'Positioning and presenting design science research for maximum impact', MIS Quarterly, Vol. 37, No. 2, pp. 337–355.
Harika, A., Balan, G., Thethi, H. P., Rana, A., Rajkumar, K. V., & Al-Allak, M. A. (2024, May). "Harnessing the power of artificial intelligence for disaster response and crisis management", In 2024 International Conference on Communication, Computer Sciences and Engineering (IC3SE) (pp. 1237-1243). IEEE.
Hevner, A.R., March, S.T., Park, J. and Ram, S. (2004) 'Design science in information systems research', MIS Quarterly, Vol. 28, No. 1, pp. 75–105.
Jaakkola, E. (2020) 'Designing conceptual articles: four approaches', AMS Review, Vol. 10, No. 1–2, pp. 18–26.
Karinshak, E. (2024). *Simulations in human-AI crisis management systems*. Journal of Contingencies and Crisis Management, 32(1), 1–29.
Li, B., Ma, J., Yin, K., Xiao, Y., Hsu, C.W. and Mostafavi, A. (2025) 'Disaster management in the era of agentic AI systems: a vision for collective human-machine intelligence for augmented resilience', arXiv preprint, arXiv:2510.16034.
Liebowitz, J. and McAlindon, P. (eds.) (2026) Leveraging AI for Business Innovation, World Scientific Publishing, Singapore.
Margherita, A., Elia, G., Solazzo, G., Gatti, L. and Poorterman, A. (2024) 'Collaborative crisis management: a coordination science framework to enhance stakeholder responses to emergencies', IEEE Transactions on Engineering Management, Vol. 72, pp. 191–201.
Nespeca, V., Comes, T., Meesters, K. and Brazier, F. (2020) 'Towards coordinated self-organization: an actor-centered framework for the design of disaster management information systems', International Journal of Disaster Risk Reduction, Vol. 51, 101887.
Okhuysen, G.A. and Bechky, B.A. (2009) 'Coordination in organizations: an integrative perspective', The Academy of Management Annals, Vol. 3, No. 1, pp. 463–502.
Otal, H.T., Stern, E. and Canbaz, M.A. (2024) 'LLM-assisted crisis management: building advanced LLM platforms for effective emergency response and public collaboration', in 2024 IEEE Conference on Artificial Intelligence (CAI), pp. 851–859.
Pantiris, G., Pallis, T. and Chountalas, P. (2025) 'Enhancing coordination and decision making in humanitarian logistics through artificial intelligence: a grounded theory approach', Logistics, Vol. 9, No. 3, 113.
Pearl, J. and Mackenzie, D. (2018) The Book of Why: The New Science of Cause and Effect, Basic Books.
Piller, C. (2024). "Generative AI and the rhetorical risks of crisis communication", Public Relations Review, 50(2), 102341.
Pimentel, A. and Veliz, C. (2025) 'Generative AI solutions for enhancing knowledge management: a literature review and research roadmap', in Proceedings of the 26th European Conference on Knowledge Management (ECKM 2025), Academic Conferences International, pp. xx–xx.
Rai, A., Constantinides, P. and Sarker, S. (2025) 'Beyond human-in-the-loop: sensemaking between artificial and human intelligence', Sustainable Futures, Vol. 7, 100216.
Reichstein, M. et al. (2025) 'Early warning of complex climate risk with integrated artificial intelligence', Nature Communications, Vol. 16, Article 2564.
Sandvik, K. B. (2025). Calibrating AI/d talk: framing perceptions, reframing policy, and deframing knowledge. Journal of International Humanitarian Action, 10(1), 11.
SAPEA (2025). Artificial Intelligence in Emergency and Crisis Management: Rapid Evidence Review Report. Munich: SAPEA. DOI: 10.5281/zenodo.17737962
Shneiderman, B. (2020). Human-centered artificial intelligence: Reliable, safe & trustworthy. International Journal of Human–Computer Interaction, 36(6), 495-504.
Turoff, M., Chumer, M., de Walle, B. V., & Yao, X. (2004). 'The design of a dynamic emergency response management information system (DERMIS)'. Journal of Information Technology Theory and Application (JITTA), Vol. 5. No. 4, 3.
Zhu, J., Ding, R., Liu, Y., & Li, M. (2024). "Large language models for coordination and cooperation: A survey", IEEE Transactions on Artificial Intelligence (in press).

Artificial Intelligence in Pharmaceutical R&D: Evidence on Trial

Mattia Vincenzo Olive, Maria Pinelli, Chiara Sgarbossa, Luca Gastaldi, Emanuele Lettieri
Politecnico di Milano, Milan (Italy)

Abstract

Pharmaceutical Research and Development (R&D) is characterised by exceptional complexity, high costs, and pervasive uncertainty, with drug development typically spanning more than a decade and requiring investments that routinely exceed several billion dollars. Artificial Intelligence (AI) has attracted growing attention as a means of improving efficiency, accelerating early prioritisation, and reducing costly late-stage failures. Yet AI integration is far from straightforward: AI outputs are probabilistic and model-based, they lack causal explanations, and their reliability depends on data quality and modelling assumptions that are difficult to validate through conventional experimental protocols. These characteristics introduce epistemic uncertainty that intersects with the stringent accountability requirements governing pharmaceutical innovation.

This paper examines how a large pharmaceutical company integrates AI into its R&D processes, focusing on the governance mechanisms that enable firms to manage AI-introduced uncertainty while meeting regulatory standards of traceability and accountability. Drawing on the Technology–Organisation–Environment (TOE) framework, we structure the analysis across three dimensions: AI output characteristics (Technology), internal decision-making and governance arrangements (Organisation), and external regulatory expectations (Environment). Data were collected through an in-depth qualitative case study comprising twelve semi-structured interviews with senior leaders, AI specialists, and R&D managers, supported by document analysis and data triangulation.

Three main findings emerge. First, the company systematically frames AI as decision support rather than decision authority: outputs are treated as signals that can focus attention and narrow options, but are not accepted as sufficient grounds for resource commitment without human review, preserving regulatory legibility and clear accountability. Second, the value attributed to AI is inseparable from documentation and traceability practices. Teams invest in decision records that connect AI inputs to subsequent validation steps, making the reasoning behind key R&D choices auditable over time. Third, AI reshapes rather than eliminates uncertainty by shifting its locus upstream in the pipeline. The company responds by embedding AI within staged decision processes defined by explicit evidence thresholds, so that the technology is most influential in early phases and progressively less decisive as development approaches regulatory-facing milestones.

Theoretically, these findings show that governance practices mediate between AI's technical potential and pharmaceutical regulation's accountability requirements, and that AI redistributes uncertainty rather than removing it, shifting its character from experimental to epistemic. Practically, the study offers guidance for firms seeking to translate AI capabilities into sustained operational value through appropriate governance infrastructure, cross-functional coordination, and clear decision responsibilities.

Keywords: Pharmaceutical R&D; Artificial Intelligence; Pharmaceutical Industry; TOE framework; Healthcare Innovation

Paper type: Academic Research Paper

1 Background

Pharmaceutical Research and Development (R&D) is widely recognised as one of the most costly, time-consuming and uncertain innovation processes in contemporary industry (DiMasi et al., 2016; Pisano, 2006; Wouters et al., 2020). On average, developing a new drug takes more than ten years and requires investments that often reach several billion dollars, while only a small share of candidate compounds eventually obtain regulatory approval (DiMasi et al., 2016; Hay et al., 2014; Schuhmacher et al., 2016). Many projects are discontinued as they move from early research to clinical testing, meaning that resources are often committed long before the final outcome is known (Kaitin and DiMasi, 2011). For this reason, early decisions about which targets, compounds and development pathways to pursue have long-term consequences on resource allocation.

In this context, Artificial Intelligence (AI) has gained increasing attention in pharmaceutical R&D. Over the past decade, AI-based systems have been introduced to support activities such as target identification, compound screening, molecular design and clinical trial planning (Mak and Pichika, 2019; Paul et al., 2021; Vamathevan et al., 2019). By processing large and diverse datasets, these systems are expected to improve early prioritisation and support earlier go or no-go decisions, with the aim of reducing costly failures at later stages of development (Schneider et al., 2020).

At the same time, AI contributes to decision-making in ways that differ from established experimental and clinical practices. AI systems typically produce rankings, scores or predictions based on statistical patterns in historical data. These outputs are probabilistic and model-based, they rarely offer causal explanations, and they cannot be validated through standard experimental protocols when they are generated (Beam and Kohane, 2018; Lipton, 2018). Their reliability depends on data quality, modelling choices and assumptions that are often difficult to fully inspect or keep stable over time (Bender et al., 2021; Leonelli and Tempini, 2020). As a result, AI does not simply speed up existing R&D activities, but also changes what counts as usable evidence at early stages of innovation.

This creates a key trade-off in pharmaceutical R&D. AI makes it possible to move prioritisation and investment decisions earlier in the development process but, at the same time, it generates new forms of uncertainty in decision-making and evidence generation. This trade-off is particularly important since pharmaceutical innovation is characterised by strong regulatory oversight and public scrutiny. To obtain marketing authorisation, firms must document how decisions were made and how evidence was produced, in forms that remain traceable and open to review by regulatory authorities over time (Abraham, 2002; Davis and Abraham, 2011). In this setting, AI-related uncertainty concerns not only accuracy, but also explainability, accountability and alignment with established scientific and regulatory standards (Topol, 2019).

Despite the growing interest in AI for pharmaceutical R&D, there is still limited empirical evidence on how companies manage this trade-off in practice. Most existing studies focus either on technical applications or on high-level ethical and regulatory discussions, offering little insight into how AI is integrated into everyday R&D processes (Floridi et al., 2018; Morley et al., 2020).

To address this gap, this working paper adopts the Technology–Organisation–Environment (TOE) (Baker, 2012; Tornatzky et al., 1990) framework as an analytical scaffold to structure the analysis across the characteristics of AI outputs (Technology), internal decision and governance arrangements (Organisation), and external expectations of accountability and reviewability (Environment).

Thus, our guiding research question is: How do technological, organisational, and environmental factors shape the integration of AI into pharmaceutical R&D, and what governance practices enable firms to translate technical potential into sustained operational and strategic value?

2 Methodology

This study is based on an in-depth qualitative case study of a large pharmaceutical company's R&D department, hereafter referred to as Company S., to examine how the uncertainty introduced by AI outputs is handled in practice.

We explore AI adoption in a real-world context, examining the strategies, challenges, and organisational dynamics of integration across development stages (Siggelkow, 2007; Yin, 2009). A single-case design provides depth and contextual specificity, capturing the regulatory, data management, and organisational barriers that condition effective implementation.

Semi-structured interviews, described by Yin (2009) as "both versatile and flexible", were used to explore how AI integration unfolds across R&D decision contexts. To ensure the validity of the study we have followed a protocol for conducting interviews, a data analysis through NVivo software and data triangulation from multiple sources (official website, internal documents and reports).

The choice of Company S. reflects the company's substantial and well-documented investments in AI across R&D functions, including collaborations with external partners spanning drug discovery, clinical trials, and operations. This breadth of AI integration makes it a suitable case for examining how the technology is adopted and governed across the full development spectrum.

Data were gathered through semi-structured interviews with 12 actors involved in AI development, data governance and R&D decision-making. The choice of participants was made to cover various roles within Company S., ensuring a holistic view of AI's organisational impact within R&D. The diversified interview sample includes senior leadership, AI specialists, and R&D managers. Over three months (February to April 2024), 12 key informants were interviewed via Zoom by at least two authors of this manuscript to increase the reliability of the notes. Each session lasted 45 to 75 minutes and was audio-recorded with prior consent. The key topics addressed during the interview included: (i) the role of AI in R&D processes; (ii) challenges and barriers encountered during implementation; (iii) ethical and regulatory issues; and (iv) organisational strategies for integrating AI into workflows. The protocol included 20 questions organised into six thematic areas: background and role, AI in drug discovery, open innovation, business model impacts, challenges and barriers, and implementation benefits.

Clean verbatim transcription ensured accurate representation of the interviews for subsequent analysis (McGrath et al., 2019). The data analysis was conducted through an inductive thematic analysis (Braun and Clarke, 2006) using NVivo 11 software, which enabled the creation of nodes to represent emerging themes, allowing for a more structured approach to data management, synthesis, and visualisation.

3 Results

3.1 Technological context

Interviewees describe AI integration as spanning multiple R&D stages, with adoption intensity varying by activity and data maturity. One informant summarised the pipeline scope and heterogeneity of Machine Learning (ML) usage as follows: "There are two major parts in the discovery… the first is drug development… and the second is, in the clinical part. Adopting ML tools varies dramatically between these areas." (Head of R&D Data and Computational Science; Interviewee 12).

In early discovery, AI is framed as a capability to explore highly complex search spaces and accelerate target identification. As the Global Head of AI/ML for R&D reported: "Starting from the very, very beginning… we call it Target ID." (Interviewee 6). By exploring vast molecular spaces, AI helps Company S. pinpoint viable drug candidates, a task otherwise unattainable without advanced algorithms. As reported by the Head of AI for Large Molecule Research, "When you look at this molecular space, it's something like 10 to the power of 60 different possibilities, and most of them are not good." (Interviewee 4). AI is thus positioned as a filtering and prioritisation layer to reduce the search space and improve precision.

Company S. integrates AI not only in early drug discovery but also in clinical activities, where AI supports cohort identification and recruitment by linking real-world information to protocol constraints. The Head of Clinical Innovation Strategy highlighted that "Looking at patient records, understanding opportunities for patient recruitment, where patients match protocol criteria is important." (Interviewee 3). This indicates that value is not limited to discovery but extends to operationally relevant clinical processes.

AI facilitates resource optimisation at Company S., enhancing both efficiency and adaptability. The Head of AI for Large Molecule Research highlighted time-saving as a critical advantage: "You can reduce the time from the target to the market, or from the target identification to the preclinical phase." (Interviewee 4). By reducing clinical trial failure rates and streamlining processes, AI improves return on investment, which is especially critical in an industry where development costs often exceed $1 billion per drug.

Data foundations emerge as the core enabling substrate. Company S.'s approach emphasises central standards and governance to enable reuse and interoperability. The Global Head of the R&D Data Office stated: "Building the data strategy, putting together the standards, data management, and data governance… allows us to build this end-to-end data ecosystem." (Interviewee 9). This technological context also aligns with a hybrid sourcing strategy: the company selects "best-in-breed" tools from multiple providers to ensure flexibility and interoperability across platforms and R&D needs.

3.2 Organisational context

Organisational readiness strongly conditions whether AI tools become embedded capabilities rather than isolated pilots. Integrating AI into resource allocation processes demands a shift in organisational culture and structure. As the Global Head of the R&D Data Office stressed: "It's the whole organisational level of upskilling that needs to happen." (Interviewee 9). This indicates that adoption is constrained not only by the availability of data scientists but by the broader ability of scientific and operational teams to interpret, trust, and appropriately use AI outputs.

AI integration also requires cross-functional routines that connect data governance, scientific validation, and business decision processes. Interviewees repeatedly framed AI as augmenting human expertise rather than displacing accountability-bearing decisions. This boundary is expressed clearly in the following statement: "AI can assist but not completely replace human decision-making processes." (Head of AI for Large Molecule Research; Interviewee 4). The implication is that workflow design must specify where AI informs decisions, how outputs are checked, and who remains responsible for final judgement.

Change management and internal communication are highlighted as central levers to reduce resistance and enable adoption across functions. One informant described an anticipatory approach: "AI in Company S.… we anticipate the forthcoming changes, and there is clear communication about the impact that these new tools will have." (Head of AI for Large Molecule Research; Interviewee 4). This suggests that scaling depends on aligning expectations, clarifying role implications, and providing time and support for workflow adaptation.

Finally, organisational structure and culture shape integration pace. In the words of one interviewee: "The pharma companies are very traditional, very hierarchical… but technology is growing fast." (Head of Molecular and Digital Histopathology; Interviewee 7). This tension, between the speed of technological change and the inertia of established structures, helps explain why even technically promising tools may face slow institutionalisation unless governance and incentives evolve accordingly.

3.3 Environmental context

The external environment affects AI integration through regulation, partnership ecosystems, and intellectual property constraints. Regulatory scrutiny is a foundational element shaping internal governance practices. The demand for explainability and traceability is both an internal trust requirement and an external compliance expectation, as indicated by the need to "explain to the FDA and the authorities exactly how you generated the data." (Head of Molecular and Digital Histopathology; Interviewee 7). This raises the bar for documentation, validation, and accountability.

A second environmental factor is the role of external collaboration as a pathway to access talent, tools, and scientific advances. Company S.'s AI strategy is described as strongly collaborative: "Collaboration is key… there is a lot that we do in-house, but much more is happening outside of Company S." (VP Global Head of Precision Medicine and Computational Biology; Interviewee 11). Another informant quantified this orientation: "At least 25 percent of our projects are in collaboration with external partners." (Executive Director of Data Science; Interviewee 10). Collaboration is nonetheless selective: the company opens specific capabilities to partners while protecting proprietary data and core scientific assets.

Third, intellectual property and patent uncertainty influences how firms share and use data with partners. The Global Head of the Data Science Organisation emphasised the need for caution: "We need to find the right balance of using external data and sharing our internal data… at least until we have defined a proper patent framework to protect our IP." (Interviewee 8). This environmental condition shapes partnership governance, data-sharing boundaries, and strategic choices about "make or buy" approaches to AI tool development.

4 Discussion and Conclusion

Across the three analytical dimensions of the TOE framework, the case reveals a consistent pattern: AI adoption in pharmaceutical R&D is primarily a governance challenge, not a technical one. The three findings below each map onto one face of this challenge — defining the boundary of AI authority, anchoring its outputs in documentation, and structuring decisions around explicit evidence thresholds — before their combined implications are drawn out. Findings show, first, that the R&D department systematically frames AI as decision support rather than decision authority. In practice, AI outputs are treated as signals that can trigger attention, open new options or help narrow down choices, but they are not accepted as sufficient grounds for committing resources on their own. Keeping formal decision rights with human actors and committees functions as a core accountability device that, on the one hand, anchors responsibility in identifiable roles and, on the other, preserves a decision rationale explainable in ways that remain compatible with regulatory expectations (Abraham, 2002; Davis and Abraham, 2011).

Second, the case highlights that the "value" of AI is closely linked to documentation practices and to the ability to reconstruct decisions over time. Because AI outputs are model-based and depend on data provenance and modelling assumptions (Bender et al., 2021; Leonelli and Tempini, 2020), teams invest in practices that make these outputs traceable and discussable, rather than treating them as black-box recommendations. The empirical material shows increased attention to "decision records" that connect AI inputs to subsequent validation steps, including why particular molecules were advanced, deprioritised or stopped in the R&D process.

Third, the analysis shows that AI reshapes uncertainty by moving it earlier and changing its form. AI supports earlier prioritisation, but this also means that critical decisions are taken at a point when experimental confirmation is still partial. Rather than removing uncertainty, AI shifts the main governance challenge from "managing failure later" to "making early decisions acceptable". The company responds by embedding AI use in staged decision processes with explicit thresholds for what counts as sufficiently robust evidence at each stage. In this arrangement, AI is most influential in the early phases, where it helps focus scarce attention, but it becomes progressively less decisive as development moves closer to regulatory-facing milestones, where evidence must be based on established experimental and clinical standards (Beam and Kohane, 2018; Topol, 2019).

The study shows that AI redistributes uncertainty over time in R&D processes and changes the kind of evidence that supports decisions in early R&D phases. The organisational decisions and arrangements to manage this process are therefore a key determinant of AI's ability to reduce R&D costs and generate societal value.

Taken together, these three mechanisms — framing AI as decision support, anchoring its value in traceability, and embedding its use within staged processes with explicit evidence thresholds — constitute the governance infrastructure through which technical potential is translated into sustained operational value. Their interdependence is significant: traceability without staged gatekeeping leaves documentation disconnected from decisions, while staged decisions without clear human accountability create gaps in regulatory legibility.

From a practical perspective, this paper contributes to understanding how pharmaceutical firms can manage operations and decision-making processes in R&D when integrating AI. Specifically, it shows that capturing value from AI requires not only technical capability, but also appropriate governance mechanisms, cross-functional coordination, and clear decision responsibilities.

Theoretically, this paper contributes to the literature on AI adoption in regulated industries in two ways. First, where existing work tends to address either technical feasibility (Paul et al., 2021; Vamathevan et al., 2019) or normative principles of explainability and accountability (Floridi et al., 2018; Morley et al., 2020), this case shows how governance practices mediate between the two — translating abstract requirements into concrete workflow arrangements. Second, while prior work has largely framed uncertainty as something AI reduces, the evidence here suggests that AI redistributes uncertainty rather than eliminates it, shifting its locus upstream and changing its character from experimental to epistemic. This distinction has implications for how firms design validation processes and structure their interactions with regulators.

This study has some limitations. First, it is based on a single qualitative case study, which provides depth but limits the generalisability of the findings across firms, therapeutic areas, and institutional contexts. Second, the analysis relies primarily on interview-based evidence collected within one company and therefore reflects managerial and organisational perceptions of AI integration more than direct performance measurement. Finally, the study captures AI adoption at a specific point in time, while many of its effects on R&D outcomes may emerge only over a longer period. The rapidly evolving regulatory environment — including the EU AI Act and emerging FDA guidance on AI-assisted drug development — and the development of foundation models with capabilities substantially different from those documented here suggest that both governance practices and their organisational implications are likely to shift. Longitudinal work would be needed to track how the arrangements observed are maintained, adapted, or replaced as the technology and its regulatory context mature.

Future research could address these limitations by examining whether different organisational models lead to different patterns of AI integration and governance. Longitudinal studies would also be valuable to understand how AI-related routines, capabilities, and trust evolve over time. In addition, further work could explore how specific governance mechanisms, such as model documentation, validation practices, and human oversight, influence the ability of firms to convert AI-based insights into accepted and sustainable R&D decisions.

5 Ethics declaration

All participants provided informed consent for audio recording and subsequent analysis. Data were anonymised to protect participant confidentiality, and the participating organisation is referred to throughout as Company S.

6 AI declaration

Artificial intelligence tools (OpenAI ChatGPT and Anthropic Claude) were employed exclusively for proofreading purposes.

References

Abraham, J. (2002). The pharmaceutical industry as a political player. The Lancet, 360(9344), 1498–1502. https://doi.org/10.1016/S0140-6736(02)11477-2

Baker, J. (2012). The Technology–Organization–Environment Framework. In Y. K. Dwivedi, M. R. Wade and S. L. Schneberger (Eds), Information Systems Theory: Explaining and Predicting Our Digital Society, Vol. 1 (pp. 231–245). Springer. https://doi.org/10.1007/978-1-4419-6108-2_12

Beam, A. L. and Kohane, I. S. (2018). Big Data and Machine Learning in Health Care. JAMA, 319(13), 1317–1318. https://doi.org/10.1001/jama.2017.18391

Bender, E. M., Gebru, T., McMillan-Major, A. and Shmitchell, S. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, FAccT '21, 610–623. https://doi.org/10.1145/3442188.3445922

Braun, V. and Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3(2), 77–101.

Davis, C. and Abraham, J. (2011). Rethinking Innovation Accounting in Pharmaceutical Regulation: A Case Study in the Deconstruction of Therapeutic Advance and Therapeutic Breakthrough. Science, Technology, and Human Values, 36(6), 791–815.

DiMasi, J. A., Grabowski, H. G. and Hansen, R. W. (2016). Innovation in the pharmaceutical industry: New estimates of R&D costs. Journal of Health Economics, 47, 20–33. https://doi.org/10.1016/j.jhealeco.2016.01.012

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P. and Vayena, E. (2018). AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations. Minds and Machines, 28(4), 689–707. https://doi.org/10.1007/s11023-018-9482-5

Hay, M., Thomas, D. W., Craighead, J. L., Economides, C. and Rosenthal, J. (2014). Clinical development success rates for investigational drugs. Nature Biotechnology, 32(1), 40–51. https://doi.org/10.1038/nbt.2786

Johnson, K. B., Wei, W., Weeraratne, D., Frisse, M. E., Misulis, K., Rhee, K., Zhao, J. and Snowdon, J. L. (2021). Precision Medicine, AI, and the Future of Personalized Health Care. Clinical and Translational Science, 14(1), 86–93. https://doi.org/10.1111/cts.12884

Kaitin, K. and DiMasi, J. (2011). Pharmaceutical Innovation in the 21st Century: New Drug Approvals in the First Decade, 2000–2009. Clinical Pharmacology and Therapeutics, 89(2), 183–188. https://doi.org/10.1038/clpt.2010.286

Leonelli, S. and Tempini, N. (Eds). (2020). Data Journeys in the Sciences. Springer International Publishing. https://doi.org/10.1007/978-3-030-37177-7

Lipton, Z. C. (2018). The mythos of model interpretability. Communications of the ACM, 61(10), 36–43. https://doi.org/10.1145/3233231

Mak, K.-K. and Pichika, M. R. (2019). Artificial intelligence in drug development: Present status and future prospects. Drug Discovery Today, 24(3), 773–780. https://doi.org/10.1016/j.drudis.2018.11.014

McGrath, C., Palmgren, P. J. and Liljedahl, M. (2019). Twelve tips for conducting qualitative research interviews. Medical Teacher, 41(9), 1002–1006. https://doi.org/10.1080/0142159X.2018.1497149

Miozza, M., Brunetta, F. and Appio, F. P. (2024). Digital transformation of the Pharmaceutical Industry: A future research agenda for management studies. Technological Forecasting and Social Change, 207.

Morley, J., Floridi, L., Kinsey, L. and Elhalal, A. (2020). From What to How: An Initial Review of Publicly Available AI Ethics Tools, Methods and Research to Translate Principles into Practices. Science and Engineering Ethics, 26(4), 2141–2168. https://doi.org/10.1007/s11948-019-00165-5

Paul, D., Sanap, G., Shenoy, S., Kalyane, D., Kalia, K. and Tekade, R. K. (2021). Artificial intelligence in drug discovery and development. Drug Discovery Today, 26(1), 80–93. https://doi.org/10.1016/j.drudis.2020.10.010

Pisano, G. P. (2006). Science Business: The Promise, the Reality, and the Future of Biotech. Harvard Business Press.

Schneider, P., Walters, W. P., Plowright, A. T., Sieroka, N., Listgarten, J., Goodnow, R. A., Fisher, J., Jansen, J. M., Duca, J. S., Rush, T. S., Zentgraf, M., Hill, J. E., Krutoholow, E., Kohler, M., Blaney, J., Funatsu, K., Luebkemann, C. and Schneider, G. (2020). Rethinking drug design in the artificial intelligence era. Nature Reviews Drug Discovery, 19(5), 353–364. https://doi.org/10.1038/s41573-019-0050-3

Schuhmacher, A., Gassmann, O. and Hinder, M. (2016). Changing R&D models in research-based pharmaceutical companies. Journal of Translational Medicine, 14(1), 105. https://doi.org/10.1186/s12967-016-0838-4

Siggelkow, N. (2007). Persuasion With Case Studies. Academy of Management Journal, 50(1), 20–24. https://doi.org/10.5465/amj.2007.24160882

Topol, E. J. (2019). High-performance medicine: The convergence of human and artificial intelligence. Nature Medicine, 25(1), 44–56. https://doi.org/10.1038/s41591-018-0300-7

Tornatzky, L. G., Fleischer, M. and Chakrabarti, A. K. (1990). The Processes of Technological Innovation. Lexington Books.

Vamathevan, J., Clark, D., Czodrowski, P., Dunham, I., Ferran, E., Lee, G., Li, B., Madabhushi, A., Shah, P., Spitzer, M. and Zhao, S. (2019). Applications of machine learning in drug discovery and development. Nature Reviews Drug Discovery, 18(6), 463–477. https://doi.org/10.1038/s41573-019-0024-5

Wouters, O. J., McKee, M. and Luyten, J. (2020). Estimated Research and Development Investment Needed to Bring a New Medicine to Market, 2009–2018. JAMA, 323(9), 844–853. https://doi.org/10.1001/jama.2020.1166

Yin, R. K. (2009). Case Study Research: Design and Methods. SAGE Publications.

HRM Strategies for Diversity and Inclusion: A Bibliometric Analysis

Roberta Romano, Simona Mormile, Gabriella Piscopo, Paola Adinolfi
University of Salerno, Fisciano, Italy

Abstract

This paper systematically examines how Human Resource Management (HRM) strategies have been employed to promote diversity and inclusion (D&I) in organizations, with the aim of identifying the main themes, trends, and strategic approaches through which inclusive workplaces are fostered. Adopting a bibliometric approach, the study analyzes a curated dataset of 870 peer-reviewed, open-access journal articles published between 2012 and 2025. Data were retrieved from Scopus and examined using the Bibliometrix package in R. The analysis combines performance analysis, country-level production analysis, trend topic analysis, and conceptual structure mapping to investigate the temporal evolution, geographical distribution, and thematic clusters characterizing the literature.

The findings reveal a steady increase in academic interest in HRM strategies for D&I, with marked acceleration from 2019 onward and a peak in 2024. This growth reflects the influence of global sociopolitical movements, regulatory initiatives, ESG pressures, and changing expectations regarding equity, representation, and inclusive work environments. Geographical analysis shows a strong concentration of publications in Anglo-speaking and industrialized countries, together with growing contributions from emerging economies. Thematic analysis identifies well-developed clusters around gender identity, ethnicity, career development, workplace psychology, and job satisfaction, while also highlighting foundational themes such as sustainability, innovation, organizational culture, gender disparity, and women's status. Emerging areas include inclusive leadership, disability inclusion, diversity in top management teams, decision-making, and artificial intelligence in HR processes. The study also identifies gaps in intersectional theorizing, systemic approaches to inclusion, and the long-term effects of D&I policies.

The value of this paper lies in its comprehensive bibliometric synthesis of the HRM and D&I literature, which maps the field's conceptual, intellectual, and social structure. Its implications are relevant for scholars, practitioners, and policymakers seeking to advance inclusive HRM practices. By uncovering the "bright side" of diversity, the study reframes D&I not as a compliance obligation but as a strategic and cultural transformation imperative.

Keywords: human resource management, diversity management, organizational strategy, bibliometric analysis, innovation.

Paper type: Academic Research Paper

1 Introduction and Literature Review

In an increasingly globalized and interconnected world, diversity and inclusion (D&I) have become strategic priorities for organizations seeking competitiveness, innovation, and social responsibility. This attention reflects demographic, social, regulatory, and institutional transformations that have reshaped organizational expectations (Mor Barak, 2015). No longer confined to compliance or corporate social responsibility, D&I are recognized as levers for innovation, employee engagement, and organizational sustainability (Shore, 2011; Roberson, 2019). Organizations are therefore required to manage diversity as a knowledge-intensive process, in which heterogeneous perspectives enhance learning, creativity, and adaptive capacity.

In organizational studies, diversity and inclusion have evolved from legal compliance frameworks toward strategic approaches shaping organizational culture and performance (Shore et al., 2018). Early workplace diversity research focused on discrimination prevention and representational equity, especially after civil rights legislation in the United States (Edelman et al., 2001). Over time, diversity has been reconceptualized as a multidimensional construct referring to visible characteristics, such as gender, ethnicity, age, and disability, and less visible dimensions, including cognitive diversity, educational background, and sexual orientation (Shore, 2011; Van Knippenberg and Mell, 2016). Inclusion refers to the extent to which individuals feel valued, respected, accepted, and able to contribute meaningfully (Nishii, 2013; Ferdman and Deane, 2014). While diversity describes workforce composition, inclusion depends on organizational practices fostering belonging, voice, and psychological safety. These dynamics are central to knowledge creation, as inclusive environments enable tacit and explicit knowledge exchange across organizational boundaries.

Human resource management (HRM) plays a central role in translating D&I principles into practices. HRM systems can either enable or constrain D&I goals (Sabharwal, 2014). Organizations may promote inclusion through recruitment and selection processes that reduce bias and widen access for underrepresented groups (Konrad, 2003), bias-aware performance management and promotion systems (Castilla, 2008), diversity training

and inclusive leadership development (Randel et al., 2018), and mentoring programs or employee resource groups supporting marginalized employees (Dobbin and Kalev, 2016). However, such initiatives should not remain isolated interventions; rather, they should be aligned with strategic HRM frameworks and embedded across the employee lifecycle (Özbilgin et al., 2011). Inclusion requires cultures that reinforce fairness, respect, and voice, making D&I a systemic rather than merely programmatic concern (Nishii, 2013; Shore et al., 2018).

Despite their relevance, HRM strategies for D&I still face significant challenges. Scholars have criticized symbolic or performative initiatives lacking structural change (Ahmed, 2012), resistance from dominant groups when inclusion threatens established hierarchies (Dobbin and Kalev, 2016), and standardized approaches that overlook contextual specificities (Tatli and Özbilgin, 2012). These critiques have encouraged context-sensitive, intersectional, and reflexive perspectives attentive to power relations, organizational politics, and marginalized employees' lived experiences (Nkomo et al., 2019). In parallel, AI adoption in HRM introduces challenges concerning the validity, transparency, and ethical use of algorithmically generated knowledge, raising concerns about bias, accountability, and governance.

Although the D&I-HRM literature has expanded, it remains fragmented across business ethics, organizational behavior, public administration, psychology, and related fields, making overarching trends, conceptual convergences, and dominant paradigms difficult to identify. Most reviews are narrative or conceptual, lacking the systematic rigor needed to map the field's evolution and structure (Zupic and Čater, 2015). Bibliometric analysis addresses these limitations by enabling quantitative assessment of temporal developments, thematic patterns, and knowledge structures (Donthu et al., 2021).

To address this gap, this paper maps academic research at the intersection of HRM and D&I, focusing on the "bright side" of diversity. It conducts a bibliometric analysis of 870 peer-reviewed, open-access journal articles published between 2012 and 2025 using Bibliometrix in R (Aria and Cuccurullo, 2017). The study asks: how scientific production on HRM strategies for D&I has evolved; which thematic areas characterize the field; and which countries have contributed most to knowledge development.

The study contributes by providing a quantitative synthesis of more than a decade of research, tracing the field's maturation and internationalization; identifying key topics and trends; and mapping the domain's conceptual and social structures to highlight research gaps and future directions, especially in underrepresented contexts and populations.

2 Methodology

This study adopts bibliometric analysis to examine the literature on promoting diversity in organizations through HRM strategies. Bibliometric analysis is a quantitative method for describing, evaluating, and monitoring academic research through bibliographic data (Garfield, 2004; White and McCain, 1998). It allows the identification of research trends, influential areas, and conceptual structures.

The analysis was conducted using Bibliometrix in R (Aria and Cuccurullo, 2017), which supports the import and analysis of bibliographic data and the construction of networks revealing conceptual, intellectual, and social structures. Biblioshiny, its web interface, further improves accessibility and usability (Moral-Muñoz *et al*., 2020). The research process included two main phases: data collection and bibliometric analysis.

3.1 Data collection

Data were retrieved from Scopus, one of the most widely used databases for scientific research (Singh *et al*., 2021). The search was conducted on September 25, 2025, using terms related to diversity and HRM. The following search string was applied to titles, abstracts, and keywords: ("Diversity management" OR "Diversity and inclusion" OR "Workplace diversity") AND ("Human resource management" OR "HRM" OR "HR strategies").

The initial query returned 12,875 documents. Several exclusion criteria were then applied. First, the analysis was limited to the period 2012–2025, reflecting the increased relevance of diversity in organizational HRM strategies from 2012 onward. Second, only English-language, published, open-access journal articles were included. Third, unrelated subject areas, such as mathematics, medicine, and energy, were excluded, while business, management and accounting, social sciences, decision sciences, and psychology were retained. After data cleaning, duplicate removal, and exclusion of irrelevant contributions, the final corpus included 870 articles uploaded to Bibliometrix in CSV format.

3.2 Data analysis

Scientific output was analyzed through annual production and country-level contributions. Annual production highlighted the temporal evolution of the topic, while country-specific production showed the geographical diffusion and relevance of the field.

Trend topic analysis was used to identify HRM strategies and diversity-related concepts over time. Keywords were extracted, and their temporal distribution was calculated. For each keyword, the median year, first quartile, and third quartile were identified. Results were visualized through a scatter plot in which each bubble represented a term, its size reflected frequency, and its position indicated the median year of use. The conceptual structure was analyzed through co-word analysis, based on the assumption that keyword co-occurrence reflects the conceptual content of documents (Callon *et al*., 1991). Keywords appearing together were connected in a network, and the Louvain clustering algorithm was used to identify thematic areas. The resulting thematic map was organized according to Callon centrality and density, distinguishing motor themes, basic themes, emerging or declining themes, and niche themes (Courtial, 1994; Coulter *et al*., 1998).

3 Findings and Discussion

4.1 Research performance

The analysis of the annual evolution of scientific production (Figure 1) reveals a steadily increasing trend, marked by several particularly significant phases. Between 2012 and 2015, the literature on the topic remained limited and marginal, with fewer than twenty publications per year. Starting from 2016, a gradual increase is observed, reflecting an initial consolidation of scholarly interest in the link between diversity management and HRM practices. A more marked turning point, however, occurs during the 2019–2021 period, when the number of publications doubles, reaching approximately 125 articles per year. This surge can be attributed to the intensification of public and academic discourse on inclusion, equity, and representation, further stimulated by global social movements and growing regulatory and reputational pressures on organizations (Özbilgin, 2019).

In particular, movements such as #MeToo have brought public attention to racial and gender inequalities, urging concrete responses from the world of work (Kirton and Greene, 2021).

Simultaneously, there has been increased visibility of advocacy for *LGBTQIA+* rights, neurodiversity, and disability inclusion, thereby expanding the very notion of diversity (Fontana, 2020; Gupta and Priyadarshi, 2020). At the regulatory level, the European Union has reinforced its commitment through initiatives such as the *Gender Equality Strategy 2020–2025* (Debusscher, 2023; Muñoz; 2022). Moreover, pressures from ESG investors, international certifications (such as *EDGE* and the *Diversity & Inclusion Index*), and the expectations of new generations of workers have made diversity a central element in talent attraction strategies (Kirton and Greene, 2021; Clementi *et al*., 2019; Ng *et al*., 2021). Until 2024, there was exponential growth in research, and 2024 marked a peak in scientific output, with almost 180 articles published, signaling a phase of greater academic attention to the role of human resources as drivers of cultural and organizational change (Qurniawan *et al*., 2024; Poulose *et al*., 2024; Ezeafulukwe *et al.*, 2024).

Data for 2025 show a decline, with a lower number of publications compared to the previous year. This decrease, however, may be due to incomplete data collection, as the year is still ongoing at the time of analysis. Overall, the observed trend suggests a gradual scientific maturation of the field, with growing attention to the management of diversity within organizations.

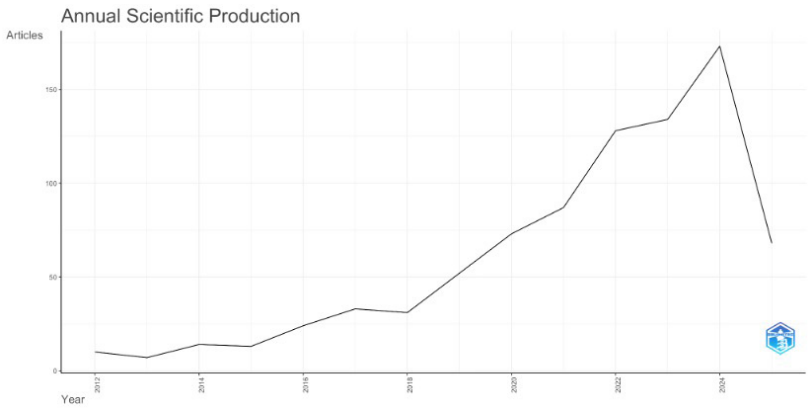


Figure 1. Annual scientific production

The geographical distribution of publications reveals a strong concentration in Anglo-speaking and industrialized countries. The United States is among the leading contributors, reflecting long-standing debates on equity, representation, and civil rights, supported by a university system strongly oriented toward empirical research (Kirton and Greene, 2021). The United Kingdom also shows substantial scientific output, consistent with a political and academic tradition that integrated diversity principles into public policies and organizational evaluation systems (Köllen, 2021).

Canada and Australia follow with significant contributions, reflecting their multicultural profiles and policies focused on ethnic, linguistic, and Indigenous inclusion (Yadav and Lenka, 2020). Among European countries, Germany, France, the Netherlands, and Italy contribute actively to diversity management research. The Italian case is developing in light of recent regulatory reforms and growing interest in the role of public administrations and firms in promoting gender equality and inclusion (Romano and Petruccioli, 2020). Law No. 162/2021, for example, introduced obligations for companies to monitor and report gender parity conditions.

Emerging countries, particularly China, India, South Korea, Brazil, Mexico, Chile, South Africa, Nigeria, and Kenya, also show growing scientific production. In these contexts, organizational diversity is often linked to economic development, institutional modernization, and corporate social responsibility (Kharroubi, 2021). Overall, the findings suggest progressive internationalization of D&I research and increasing recognition of diversity as a strategic factor for organizational competitiveness and sustainability.

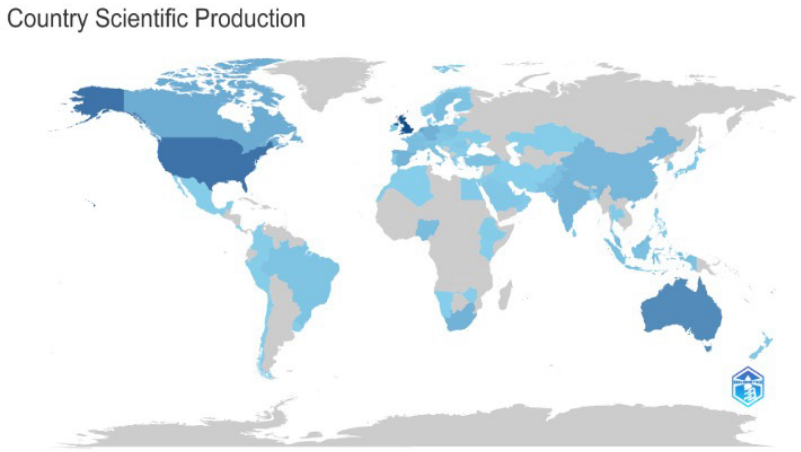


Figure 2. Country scientific production

4.2 Key themes and topics

Trend topic analysis reveals the evolution of research interests over time and indicates increasing theoretical sophistication. Recent and recurrent topics focus on the intersection between diversity and new work dynamics, especially leadership, workplace, employees, working conditions, and artificial intelligence, which have gained relevance since 2020. These terms suggest a broadening of diversity management toward employee well-being, hybrid work, and digitalization (Köllen, 2022; Dwivedi, 2025; Kamasak *et al.*, 2023).

The presence of terms such as female, male, employment, and middle-aged highlights attention to inequalities related to gender, age, and labor market access. Gender-related research appears to be moving from binary categories toward broader reflections on historically marginalized groups, including LGBTQIA+ individuals and people with migratory backgrounds, although the latter remain less prominent (Fontana, 2020; Köllen, 2021).

The growing relevance of perception indicates greater attention to qualitative and organizational-psychological approaches, focused on employees' subjective experiences of inclusive policies and practices (Yadav and Lenka, 2020). Similarly, terms related to sustainability and learning show the expansion of the field toward systemic and long-term approaches linking social equity and sustainable performance (Köllen, 2022; Dzhengiz, 2020). This development aligns with ESG criteria and governance transformations that encourage organizations to make diversity commitments visible as part of their strategic identity.

Organizational culture and cultural diversity remain central, confirming that effective HR practices depend not only on formal policies but also on cultural assumptions that are often difficult to transform (Ince, 2023; Park, 2020). The emergence of artificial intelligence is particularly relevant, as it marks growing scholarly interest in how digital technologies may promote or distort organizational diversity. Research is beginning to examine

whether algorithms used in recruitment, evaluation, and development processes introduce bias or support more equitable and transparent criteria (Rathore *et al*., 2022; Shams *et al*., 2025).

Finally, terms such as racism, satisfaction, career, and attitude indicate that the literature has also explored connections between diversity and job satisfaction, career aspirations, attitudes toward inclusion, and systemic inequalities, although these areas remain less developed (Van Knippenberg *et al.*, 2020; Maji *et al*., 2024).

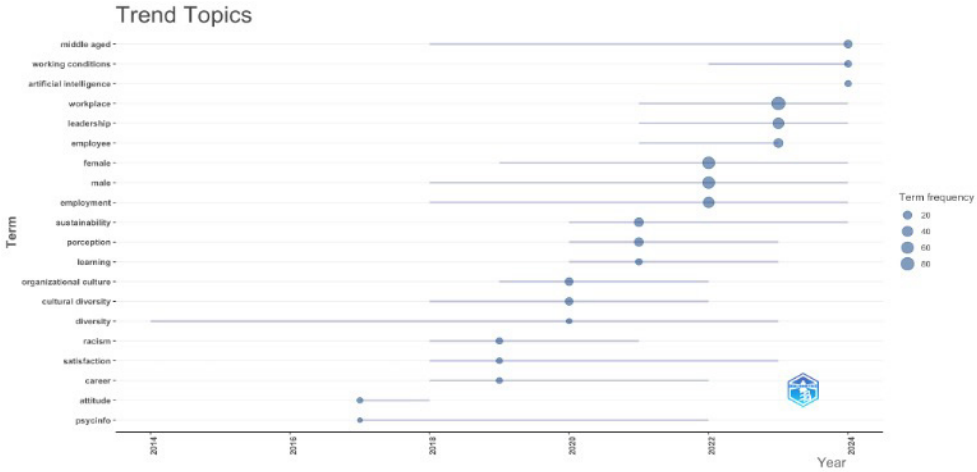


Figure 3. Trend topics analysis

The thematic map provides a visual representation of the conceptual structure of the literature, organizing themes according to development and centrality. Motor themes, located in the upper-right quadrant, are the most developed and central. These include ethnicity, gender identity, and career, which represent core pillars of the academic debate and indicate consolidated interest in identity dynamics and career trajectories in relation to diversity policies.

From an HRM perspective, these findings suggest the importance of ethnic and cultural inclusion, equitable career development, and recognition of gender identities (Morfaki and Morfaki, 2024; Yadav and Lenka, 2020). Relevant practices include mentoring for underrepresented groups and inclusive career pathways (Van Knippenberg *et al*., 2020; Williams *et al.*, 2023). The same area includes workplace, psychology, and job satisfaction, pointing to an integrated approach between inclusive HRM and psychosocial well-being (Köllen, 2022). HR strategies may therefore involve psychologically safe environments, employee support services, and systematic monitoring of organizational climate (Dwivedi, 2025; Kamasak *et al*., 2023).

Basic themes, positioned at the center of the map, are essential but still developing. The cluster composed of sustainability, innovation, and organizational framework represents the theoretical-practical infrastructure of the field, showing increasing attention to sustainability and social innovation as enabling contexts for inclusion. HRM strategies in this area may include diversity measurement in sustainability reports, integration of inclusive goals into organizational missions, and cross-functional task forces for social innovation (Walker *et al*., 2023; Naciti *et al.*, 2022).

Another central cluster includes organizational culture, gender disparity, and women's status, revealing that structural inequalities remain critical. HR practices should therefore transform organizational culture through targeted training, revision of decision-making processes, and promotion of female leadership across organizational levels (Triana *et al*., 2021; Aktepe, 2020).

Emerging or declining themes, located in the lower-left quadrant, include leadership, perception, decision-making, disability, and top management teams. Although currently less developed or marginal, these themes have significant potential, especially from a strategic HRM perspective. They represent opportunities for future research on inclusive leadership, representation at senior levels, disability inclusion, and diversity criteria in executive decision-making (Ince, 2023; Ezeafulukwe, 2024). These areas have important implications for HR innovation and organizational transformation.

Finally, niche themes in the upper-left quadrant include minority group, racism, and emotion. These themes are internally developed but less central to the broader discourse. Nevertheless, they raise important questions concerning systemic inequalities and the emotional dimensions of inclusion. HRM strategies may include training on emotional intelligence and implicit bias, listening tools for organizational minorities, and anti-discrimination policies that are experienced as credible and authentic by employees (Ogbonna, 2025; Nweiser and Dajnoki, 2022).

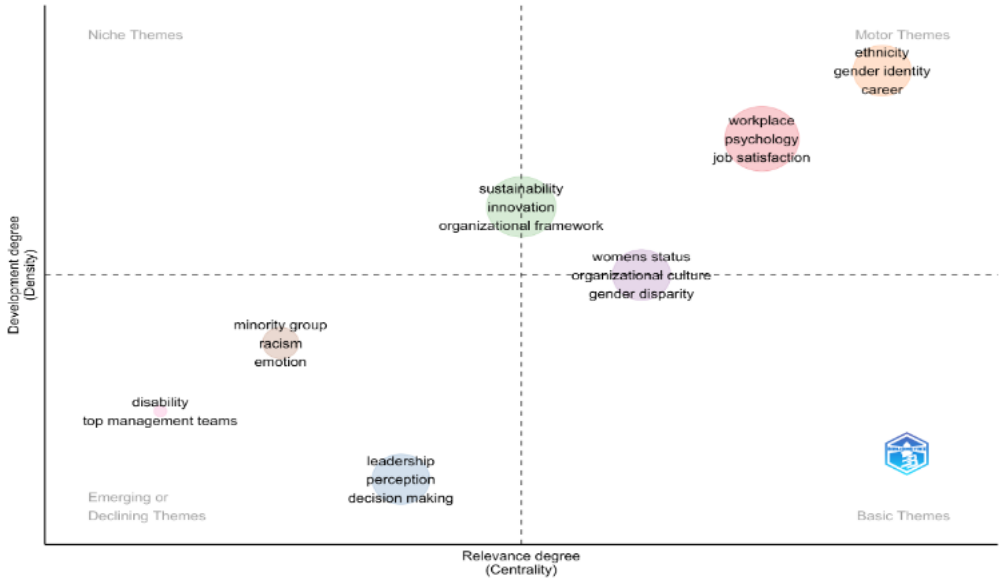


Figure 4. Thematic analysis

4 Conclusions

This study explored how the scientific literature has addressed the promotion of diversity in organizations through HRM strategies by conducting a bibliometric analysis of 870 articles published between 2012 and 2025. The aim was to identify key trends, conceptual developments, and the geographical distribution of publications, offering a systematic overview of the field.

The findings show a steadily growing body of research, with a significant acceleration from 2019 onward. This growth coincides with intensified public debates, regulatory pressures, and organizational attention to inclusion. Trend topic analysis highlights the emergence of inclusive leadership, workplace well-being, artificial intelligence, and working conditions, suggesting that diversity management is increasingly intertwined with technological, cultural, and organizational transformations. These developments indicate a shift toward hybrid knowledge ecosystems, where human judgment, relational intelligence, and AI-driven analytical capabilities may jointly support inclusive and sustainable decision-making.

The thematic map distinguishes established areas, such as gender identity and ethnicity; foundational but developing themes, such as sustainability and organizational culture; and promising emerging areas, including disability, inclusive decision-making, and diversity within top management teams.

Theoretically, the study contributes to mapping the D&I-HRM field, highlighting the transition from normative and compliance-oriented approaches to more systemic, intersectional, and strategic perspectives. However, relevant gaps remain. The literature still lacks an integrated theorization of diversity as a driver of organizational and social innovation. Further development is also needed on the relationship between diversity, organizational power structures, and the long-term effects of inclusion policies. Intersectional approaches remain underexamined, while the symbolic role of diversity in shaping organizational identity is still marginal.

Practically, the findings suggest several implications for HR professionals, public institutions, and organizational leaders. Organizations should implement structured mentoring programs for underrepresented employees, with clear objectives, accountability mechanisms, and career advancement pathways. Such programs may also serve as mechanisms for transferring tacit knowledge, strengthening relational learning, and supporting inclusive knowledge-sharing processes. D&I objectives should also be embedded in corporate missions, strategic plans, measurable KPIs, and executive evaluations. Moreover, robust D&I reporting frameworks are needed to monitor gender and ethnicity representation, pay equity, promotion rates, and hierarchical distribution.

Training on unconscious bias, inclusive behaviors, and emotional intelligence should be implemented as part of continuous learning strategies rather than isolated initiatives. Inclusive leadership models should move beyond symbolic representation and involve investment in cultural intelligence, empathic communication, and participatory decision-making. Diverse leadership should also be reflected in boards, C-suite positions, and key decision-making bodies.

Finally, organizations should critically assess artificial intelligence in HRM, especially in recruitment, performance evaluation, and promotion. AI tools may reduce human bias and enhance transparency when

properly designed, but they can also reproduce inequalities if trained on biased historical data. Their development should therefore follow inclusive design principles, involve cross-functional oversight, and include periodic bias audits to ensure fairness and accountability.

This study has limitations. The dataset is based only on Scopus and English-language publications, which may underrepresent relevant contributions from other databases and non-English-speaking contexts. Future research should expand the dataset to include sources such as Web of Science and non-English publications. Qualitative and mixed-method studies could also deepen understanding of the meanings and practices associated with diversity. Longitudinal research should examine the long-term impact of inclusive HRM strategies, while comparative studies across sectors and national contexts could clarify how cultural, institutional, and regulatory conditions influence policy effectiveness. Finally, marginal topics such as neurodiversity, algorithmic decision-making, and the emotional, symbolic, and value-laden dimensions of diversity should be explored more systematically. Advancing these directions may consolidate a transformative approach to diversity management, improving organizational relationships, internal equity, and the social sustainability of work.

5 Ethics Declaration

This study did not require ethical clearance, as it is based exclusively on bibliometric analysis of publicly available, peer-reviewed journal articles. No human participants, personal data, interviews, surveys, experiments, or sensitive information were involved in the research process.

6 AI Declaration

AI tools, specifically ChatGPT, were used exclusively for language editing, improving clarity, and refining the academic style of the manuscript.

References

Ahmed, S. (2012) On Being Included: Racism and Diversity in Institutional Life, Duke University Press.

Aktepe, S.O. (2020) “Gender Inequality in Work Organizations: What HRM Practices Mean for Gender Inequality”, Macro and Micro-Level Issues Surrounding Women in the Workforce: Emerging Research and Opportunities, pp 1-36.

Aria, M. and Cuccurullo, C. (2017) “bibliometrix: An R-tool for Comprehensive Science Mapping Analysis”, Journal of Informetrics, Vol. 11, No. 4, pp 959-975.

Callon, M., Courtial, J.P. and Laville, F. (1991) “Co-word Analysis as a Tool for Describing the Network of Interactions between Basic and Technological Research: The Case of Polymer Chemistry”, Scientometrics, Vol. 22, No. 1, pp 155-205.

Castilla, E.J. (2008) “Gender, Race, and Meritocracy in Organizational Careers”, American Journal of Sociology, Vol. 113, No. 6, pp 1479-1526.

Clementi, L., Cavadini, R., Casasopra, F., Rabozzi, M., Notargiacomo, S. and Santambrogio, M.D. (2019) “Diversity and Inclusion: Buzzword or Real Value?”, Paper read at 2019 IEEE Global Engineering Education Conference (EDUCON), April, pp 1228-1236.

Coulter, N., Monarch, I. and Konda, S. (1998) “Software Engineering as Seen through Its Research Literature: A Study in Co-word Analysis”, Journal of the American Society for Information Science, Vol. 49, No. 13, pp 1206-1223.

Courtial, J. (1994) “A Co-word Analysis of Scientometrics”, Scientometrics, Vol. 31, No. 3, pp 251-260.

Debusscher, P. (2023) “The EU Gender Equality Strategy 2020-2025: The Beginning of a New Season?”, in Social Policy in the European Union: State of Play 2022: Policymaking in a Permacrisis, ETUI; OSE, pp 91-110.

Dobbin, F. and Kalev, A. (2016) “Why Diversity Programs Fail”, Harvard Business Review, Vol. 94, No. 7, p 14.

Donthu, N., Kumar, S., Pattnaik, D. and Lim, W.M. (2021) “A Bibliometric Retrospection of Marketing from the Lens of Psychology: Insights from Psychology & Marketing”, Psychology & Marketing, Vol. 38, No. 5, pp 834-865.

Dwivedi, D. (2025) “Emotional Intelligence and Artificial Intelligence Integration Strategies for Leadership Excellence”, Advances in Research, Vol. 26, No. 1, pp 84-94.

Dzhengiz, T. (2020) “A Literature Review of Inter-organizational Sustainability Learning”, Sustainability, Vol. 12, No. 12, p 4876.

Edelman, L.B., Fuller, S.R. and Mara-Drita, I. (2001) “Diversity Rhetoric and the Managerialization of Law”, American Journal of Sociology, Vol. 106, No. 6, pp 1589-1641.

Ezeafulukwe, C., Onyekwelu, S.C., Onyekwelu, N.P., Ike, C.U., Bello, B.G. and Asuzu, F.O. (2024) “Best Practices in Human Resources for Inclusive Employment: An In-depth Review”, International Journal of Science and Research Archive, Vol. 11, No. 1, pp 1286-1293.

Ferdman, B.M. and Deane, B. (2014) Diversity at Work: The Practice of Inclusion.

Fontana, E. (2020) “Managing Diversity through Transgender Inclusion in Developing Countries: A Collaborative Corporate Social Responsibility Initiative from Bangladesh”, Corporate Social Responsibility and Environmental Management, Vol. 27, No. 6, pp 2548-2562.

Garfield, E. (2004) “Historiographic Mapping of Knowledge Domains Literature”, Journal of Information Science, Vol. 30, No. 2, pp 119-145.
Ince, F. (2023) “Transformational Leadership in a Diverse and Inclusive Organizational Culture”, Handbook of Research on Promoting an Inclusive Organizational Culture for Entrepreneurial Sustainability, pp 188-201.
Kamasak, R., Alkan, D.P. and Yalcinkaya, B. (2023) “Emerging Trends of Industry 4.0 in Equality, Diversity, and Inclusion Implementations”, Contemporary Approaches in Equality, Diversity and Inclusion: Strategic and Technological Perspectives, pp 129-148.
Kharroubi, D. (2021) “Global Workforce Diversity Management: Challenges across the World”, Ekonomicko-manazerske Spektrum, Vol. 15, No. 1, pp 28-37.
Kirton, G. and Greene, A.M. (2021) The Dynamics of Managing Diversity and Inclusion: A Critical Approach, Routledge.
Köllen, T. (2021) “Diversity Management: A Critical Review and Agenda for the Future”, Journal of Management Inquiry, Vol. 30, No. 3, pp 259-272.
Köllen, T. (2022) “What Makes a Good Diversity Manager?: A Virtue-based Perspective”, Business and Professional Ethics Journal.
Konrad, A.M. (2003) “Special Issue Introduction: Defining the Domain of Workplace Diversity Scholarship”, Group & Organization Management, Vol. 28, No. 1, pp 4-17.
Maji, S., Yadav, N. and Gupta, P. (2024) “LGBTQ+ in Workplace: A Systematic Review and Reconsideration”, Equality, Diversity and Inclusion: An International Journal, Vol. 43, No. 2, pp 313-360.
Mor Barak, M.E. (2015) “Inclusion Is the Key to Diversity Management, but What Is Inclusion?”, Human Service Organizations: Management, Leadership & Governance, Vol. 39, No. 2, pp 83-88.
Moral-Muñoz, J.A., Herrera-Viedma, E., Santisteban-Espejo, A. and Cobo, M.J. (2020) “Software Tools for Conducting Bibliometric Analysis in Science: An Up-to-date Review”, Profesional de la Información, Vol. 29, No. 1.
Morfaki, C. and Morfaki, A. (2022) “Managing Workforce Diversity and Inclusion: A Critical Review and Future Directions”, International Journal of Organizational Leadership, Vol. 11, No. 4.
Muñoz, I.G. (2022) “Critical Analysis of the Role of the European Union on Gender Equality Progress. The Case of Research Policy”, Revista de Estudios Europeos, No. 79, pp 457-472.
Naciti, V., Cesaroni, F. and Pulejo, L. (2022) “Corporate Governance and Sustainability: A Review of the Existing Literature”, Journal of Management and Governance, Vol. 26, No. 1, pp 55-74.
Ng, E.S., Stamper, C.L. and Klarsfeld, A. (2021) “Introduction to the Handbook: A Review of 23 Diversity and Inclusion Indices”, Handbook on Diversity and Inclusion Indices, pp 1-15.
Nishii, L.H. (2013) “The Benefits of Climate for Inclusion for Gender-diverse Groups”, Academy of Management Journal, Vol. 56, No. 6, pp 1754-1774.
Nkomo, S.M., Bell, M.P., Roberts, L.M., Joshi, A. and Thatcher, S.M. (2019) “Diversity at a Critical Juncture: New Theories for a Complex Phenomenon”, Academy of Management Review, Vol. 44, No. 3, pp 498-517.
Nweiser, M.T.A. and Dajnoki, K. (2022) “The Importance of Diversity Management in Relation with Other Functions of Human Resource Management: A Systematic Review”.
Ogbonna, E. (2025) “Reflections on Achieving Anti-racism in Organisations: The Role of Human Resource Management Scholars and Practitioners”, Human Resource Management Journal, Vol. 35, No. 2, pp 428-453.
Özbilgin, M.F. (2019) Managing Diversity and Inclusion in the Global Value Chain.
Özbilgin, M.F., Beauregard, T.A., Tatli, A. and Bell, M.P. (2011) “Work-life, Diversity and Intersectionality: A Critical Review and Research Agenda”, International Journal of Management Reviews, Vol. 13, No. 2, pp 177-198.
Park, J.S.Y. (2020) “Translating Culture in the Global Workplace: Language, Communication, and Diversity Management”, Applied Linguistics, Vol. 41, No. 1, pp 109-128.
Poulose, S., Bhattacharjee, B. and Chakravorty, A. (2024) “Determinants and Drivers of Change for Digital Transformation and Digitalization in Human Resource Management: A Systematic Literature Review and Conceptual Framework Building”, Management Review Quarterly, pp 1-26.
Qurniawan, H., Saragihi, I.S. and Suhendro, D. (2024) “Harnessing Diversity: The Role of Inclusive HR Practices in Driving Innovation and Organizational Growth”, INVEST: Jurnal Inovasi Bisnis dan Akuntansi, Vol. 5, No. 1, pp 233-239.
Randel, A.E., Galvin, B.M., Shore, L.M., Ehrhart, K.H., Chung, B.G., Dean, M.A. and Kedharnath, U. (2018) “Inclusive Leadership: Realizing Positive Outcomes through Belongingness and Being Valued for Uniqueness”, Human Resource Management Review, Vol. 28, No. 2, pp 190-203.
Rathore, B., Mathur, M. and Solanki, S. (2022) “An Exploratory Study on Role of Artificial Intelligence in Overcoming Biases to Promote Diversity and Inclusion Practices”, Impact of Artificial Intelligence on Organizational Transformation, pp 147-164.
Roberson, Q.M. (2019) “Diversity in the Workplace: A Review, Synthesis, and Future Research Agenda”, Annual Review of Organizational Psychology and Organizational Behavior, Vol. 6, No. 1, pp 69-88.
Romano, A. and Petruccioli, R. (2020) “Gender Diversity Management, Culture Inclusive e Sfide dell’Attualità. Una Review Sistematica della Letteratura”, Education Sciences & Society-Open Access, Vol. 11, No. 1.
Sabharwal, M. (2014) “Is Diversity Management Sufficient? Organizational Inclusion to Further Performance”, Public Personnel Management, Vol. 43, No. 2, pp 197-217.
Shams, R.A., Zowghi, D. and Bano, M. (2025) “AI and the Quest for Diversity and Inclusion: A Systematic Literature Review”, AI and Ethics, Vol. 5, No. 1, pp 411-438.

Shore, L.M., Cleveland, J.N. and Sanchez, D. (2018) “Inclusive Workplaces: A Review and Model”, Human Resource Management Review, Vol. 28, No. 2, pp 176-189.

Shore, L.M., Randel, A.E., Chung, B.G., Dean, M.A., Holcombe Ehrhart, K. and Singh, G. (2011) “Inclusion and Diversity in Work Groups: A Review and Model for Future Research”, Journal of Management, Vol. 37, No. 4, pp 1262-1289.

Singh, V.K., Singh, P., Karmakar, M., Leta, J. and Mayr, P. (2021) “The Journal Coverage of Web of Science, Scopus and Dimensions: A Comparative Analysis”, Scientometrics, Vol. 126, pp 5113-5142.

Tatli, A. and Özbilgin, M.F. (2012) “An Emic Approach to Intersectional Study of Diversity at Work: A Bourdieuan Framing”, International Journal of Management Reviews, Vol. 14, No. 2, pp 180-200.

Triana, M.D.C., Gu, P., Chapa, O., Richard, O. and Colella, A. (2021) “Sixty Years of Discrimination and Diversity Research in Human Resource Management: A Review with Suggestions for Future Research Directions”, Human Resource Management, Vol. 60, No. 1, pp 145-204.

Van Knippenberg, D. and Mell, J.N. (2016) “Past, Present, and Potential Future of Team Diversity Research: From Compositional Diversity to Emergent Diversity”, Organizational Behavior and Human Decision Processes, Vol. 136, pp 135-145.

Van Knippenberg, D., Nishii, L.H. and Dwertmann, D.J.G. (2020) “Synergy from Diversity: Managing Team Diversity to Enhance Performance”, Behavioral Science & Policy, Vol. 6, No. 1, pp 75-92.

Walker, B., Crépin, A.S., Nyström, M., Anderies, J.M., Andersson, E., Elmqvist, T. and Vincent, J.R. (2023) “Response Diversity as a Sustainability Strategy”, Nature Sustainability, Vol. 6, No. 6, pp 621-629.

White, H.D. and McCain, K.W. (1998) “Visualizing a Discipline: An Author Co-citation Analysis of Information Science, 1972–1995”, Journal of the American Society for Information Science, Vol. 49, No. 4, pp 327-355.

Williams, J.S., Walker, R.J., Burgess, K.M., Shay, L.A., Schmidt, S., Tsevat, J. and Egede, L.E. (2023) “Mentoring Strategies to Support Diversity in Research-focused Junior Faculty: A Scoping Review”, Journal of Clinical and Translational Science, Vol. 7, No. 1, e21.

Yadav, S. and Lenka, U. (2020) “Diversity Management: A Systematic Review”, Equality, Diversity and Inclusion: An International Journal, Vol. 39, No. 8, pp 901-929.

Zupic, I. and Čater, T. (2015) “Bibliometric Methods in Management and Organization”, Organizational Research Methods, Vol. 18, No. 3, pp 429-472.

From Quantitative ESG Data to Qualitative Social Insights: A Low-Code AI Agent for Decision Support

Riccardo Censi[1], Paola Campana[1], Fulvio Schettino[1], Chiara De Pucchio[1], Valentina Ferreri[2]
[1]Sapienza University of Rome, Rome, Italy
[2]Luiss Guido Carli University, Rome, Italy

Abstract

This paper explores the use of Large Language Model (LLM)-based agents as tools for supporting knowledge creation from Environmental, Social and Governance (ESG) data, with a specific focus on the social dimension. Despite the growing availability of ESG data, organisations still face substantial challenges in transforming quantitative indicators into actionable insights that can effectively support strategic decision-making. Current reporting tools mainly provide descriptive outputs, while more advanced AI-based solutions often require technical expertise that limits their practical adoption. This study addresses that gap by proposing an accessible, low-cost decision support system built on open-source technologies. The research adopts a Design Science Research (DSR) approach and develops a conversational ESG advisory agent built on Langflow, an open-source low-code platform, and Llama 4 Scout 17B, a large language model served free of charge via Groq. The agent receives structured ESG indicators through a guided conversational interface, computes relevant metrics, compares them against sector benchmarks derived from principal component analysis and cluster analysis of real energy-sector data, and produces a qualitative assessment report including an improvement roadmap aligned with Global Reporting Initiative (GRI) 401-406, European Sustainability Reporting Standards (ESRS) S1-S2, and Sustainable Development Goals (SDGs) 3, 5, and 10. Evaluation on a synthetic dataset of 40 fictitious energy companies shows that the agent achieves 98.1% value accuracy and 93.2% judgment accuracy on eight of the eleven variables tested. The remaining three variables, specifically the Gender Pay Gap and the two Cultural Diversity indicators, show significantly lower accuracy ranging from 33% to 45%, attributable to the model's limited capacity to handle competing instructions and maintain parallel reasoning tracks. Through iterative DSR testing, the study finds that restructuring the input to eliminate ambiguity is a more reliable corrective strategy than refining the prompt, suggesting a broader design principle for low-code AI systems based on mid-sized models. The paper contributes to the emerging literature on AI-driven sustainability analytics by proposing and evaluating a replicable, low-cost system that bridges the gap between quantitative ESG reporting and qualitative organisational insight. From a practical standpoint, the findings demonstrate that accessible AI architectures can support meaningful ESG knowledge creation even in resource-constrained organisational settings, acting as a cognitive support tool for human decision-makers rather than replacing them.

Keywords: Hybrid Knowledge creation; ESG performance; Artificial Intelligence; Decision support system; Social sustainability

Paper type: Academic Research Paper

1 Introduction

In recent years, Environmental, Social and Governance (ESG) criteria have become a key reference for evaluating the sustainability and responsibility of organisations. Their relevance has increased significantly due to growing pressure from investors, regulators, and a wide range of stakeholders (Sierdovski et al., 2022; Censi et al., 2025). As a result, companies are producing an increasing amount of ESG-related data, supported by the diffusion of reporting standards and digital tools for data collection and analysis.

Despite this growth in data availability, organisations still face substantial challenges in transforming ESG data into actionable knowledge that can effectively support strategic decision-making. This issue is particularly evident in the social dimension of ESG, which includes complex and multidimensional aspects such as employee well-being, diversity and inclusion, workplace safety, and stakeholder relationships (Cui, 2024; Govindan et al., 2025). These elements are difficult to capture through purely quantitative indicators, as they often require contextual interpretation and qualitative assessment.

Current ESG reporting practices and analytics tools mainly focus on numerical indicators, aggregated scores, and descriptive dashboards (Nielsen, 2023). While these tools are useful for monitoring performance, they provide limited support for understanding causal relationships, identifying critical variables, and defining improvement strategies. At the same time, more advanced solutions based on Artificial Intelligence (AI) often require high levels of technical expertise, which reduces their accessibility and limits their adoption within organisations (Mustafa et al., 2025). In this context, recent developments in Large Language Models (LLMs), a

class of AI systems capable of processing and generating natural language, open new opportunities for enhancing the interpretation of ESG data. These models can support the transformation of structured data into qualitative insights, contributing to knowledge creation processes and organisational sensemaking (Rahmani et al., 2025; Engstrom et al., 2024). However, their integration into decision support systems remains an emerging area, especially in relation to accessible and low-code solutions.

Low-code approaches refer to development environments that allow users to design and implement applications with minimal programming effort. These approaches can reduce barriers to the adoption of AI technologies, enabling a broader range of users to interact with advanced analytical tools. In the ESG domain, combining low-code platforms with LLMs may provide a practical way to bridge the gap between quantitative data and qualitative interpretation. Based on these considerations, this study aims to explore the use of LLM-based agents as tools for supporting knowledge creation from ESG data, with a specific focus on the social dimension. The research adopts a Design Science Research (DSR) approach (Verschuren & Hartog, 2005; Delport et al., 2024) and develops a technological artefact capable of receiving structured ESG indicators as input and producing qualitative assessments and strategic recommendations as output. The proposed system is designed using a low-code architecture, integrating data analytics and language models within a single workflow.

The paper contributes to both theory and practice. From a theoretical perspective, it extends existing studies on the role of AI in knowledge creation and organisational decision-making, particularly in the context of sustainability (Gandia et al., 2025). From a practical perspective, it proposes a replicable and accessible model for building decision support systems that go beyond descriptive reporting and support a more interpretative and strategic use of ESG data.

2 Theoretical Background

The increasing relevance of ESG criteria has stimulated a growing body of research on how organisations collect, analyse, and use sustainability-related data. Three main streams of literature are particularly relevant: ESG measurement and social sustainability; knowledge creation and organisational sensemaking; and the role of AI in decision support systems.

2.1 ESG measurement and social sustainability

ESG frameworks are widely used to assess corporate sustainability performance, integrating environmental, social, and governance dimensions into a unified perspective (Sierdovski et al., 2022). Among these, the social dimension presents specific challenges due to its qualitative and context-dependent nature. It includes aspects such as employee well-being, diversity and inclusion, labour conditions, and relationships with stakeholders and communities (Govindan et al., 2025). While the availability of ESG data has increased significantly, much of the existing literature highlights the limitations of current measurement practices. ESG reporting systems often rely on standardised indicators and aggregated scores, which facilitate comparison across firms but may oversimplify complex social phenomena (Nielsen, 2023). As a result, important nuances and contextual factors are often lost, limiting the ability of organisations to derive meaningful insights from the data. Recent studies have also emphasised the role of digital transformation in improving ESG performance, suggesting that data-driven approaches can support more effective sustainability strategies (Cui, 2024). However, the challenge remains not only to collect and analyse data, but also to interpret it in a way that supports decision-making, especially when dealing with multidimensional and interrelated variables.

2.2 Knowledge creation and organisational sensemaking

The difficulty in interpreting ESG data can be understood through the lens of knowledge creation and sensemaking. Knowledge creation refers to the process through which data and information are transformed into meaningful insights that can guide action. In organisational contexts, this process often involves interpreting complex and ambiguous information, identifying patterns, and constructing shared understanding. Recent research has highlighted how AI technologies can support these processes by assisting in the interpretation and contextualisation of data (Engstrom et al., 2024). In particular, AI systems can contribute to sensemaking by generating narratives, identifying relationships between variables, and offering alternative perspectives on organisational performance. The role of AI is not limited to prediction or automation. Instead, it can act as a cognitive support tool that enhances human decision-making by providing structured interpretations and suggestions (Qin, 2024). This perspective is particularly relevant for ESG analysis, where decision-makers must deal with incomplete information, qualitative judgements, and multiple stakeholder perspectives.

2.3 Artificial Intelligence and low-code decision support systems

The application of AI to ESG analytics has attracted increasing attention in recent years. Studies have shown that AI can improve the quality, reliability, and relevance of non-financial reporting, supporting both internal decision-making and external communication (Majekodunmi, 2025; Mustafa et al., 2025). However, many AI-based solutions remain complex and require specialised technical skills, limiting their practical adoption.

LLMs represent a recent development in AI that can address some of these limitations. LLMs are capable of processing both structured and unstructured data and generating natural language outputs, making them suitable for tasks that involve interpretation and explanation (Rahmani et al., 2025). Their ability to translate quantitative inputs into qualitative insights makes them particularly relevant for ESG contexts. At the same time, the emergence of low-code platforms has made it easier to design and deploy AI-based applications. Low-code environments allow users to build data pipelines and integrate different components, such as data analytics tools and language models, with limited programming effort. This reduces barriers to entry and enables a wider range of organisations to experiment with AI-driven solutions. From a decision support perspective, this combination aligns with the concept of Decision Support Systems (DSS), which are designed to assist rather than replace human decision-makers. Early work on DSS highlights the importance of integrating data, models, and user interaction in a coherent system that supports problem understanding and solution development (Poch et al., 2004). Building on these insights, recent research has started to explore the role of AI as a tool for enhancing knowledge and supporting sustainability-oriented decision-making (Gandia et al., 2025). The integration of LLMs within low-code architectures can thus be seen as a promising approach to bridge the gap between data availability and knowledge creation, particularly in complex domains such as social sustainability.

3 Methodology

3.1 Data and empirical setting

The empirical foundation of the agent is based on data collected from the LPSG dataset, focusing on the energy sector over the period 2020-2024. The dataset includes 16 variables for a total of 66 firms and is used to derive sector benchmarks and support the analytical components embedded in the agent.

A thorough data cleaning and preprocessing phase was conducted to ensure suitability for econometric modeling. Out of the initial 16 variables, 9 were retained based on data completeness and temporal consistency, ensuring sufficient observations across the four-year period.

The correlation analysis reveals a generally low level of dependence among variables, suggesting limited multicollinearity within the dataset. The only notable exception is the strong correlation between the Social Pillar Score and the ESG Score (0.82), which is expected given that the social dimension is a core component of the overall ESG metric.

From an economic perspective, several meaningful relationships emerge. Firm size, proxied by the number of employees, is positively associated with both ESG performance and the Social Pillar Score, indicating that larger firms tend to adopt more advanced sustainability practices. Furthermore, governance and diversity-related variables exhibit moderate correlations with ESG performance, highlighting a consistent relationship between the quality of governance structures and attention to sustainability issues.

In contrast, financial variables such as Return on Equity (ROE) and dividend yield show weak correlations with ESG indicators, suggesting the absence of a strong direct association within the analyzed sample. Similarly, the time variable does not display significant trends, indicating that the observed differences are primarily cross-sectional rather than driven by temporal dynamics.

The Principal Component Analysis (PCA) indicates that the variables contributing most to total variance (28%) are the number of employees, board diversity, and executive members' gender diversity.

Overall, these findings define the empirical structure underpinning the agent, informing both the construction of sector benchmarks and the representation of the decision environment in which the agent operates.

3.2 Agent architecture

The agent runs on Langflow, an open-source platform that lets users assemble LLM-based pipelines visually, without writing code (Taulli & Deshmukh, 2025). Four nodes make up the pipeline: (1) a Chat Input node that captures the user's messages; (2) a Prompt Template node where all domain knowledge resides, variable definitions, formulas, sector benchmarks, PCA results, and cluster profiles; (3) a Chat Memory node that tracks

conversation history across multiple turns; and (4) an LLM node that sends the assembled prompt to the language model and returns the response.

The model selected is Llama 4 Scout 17B, served free of charge via Groq. The choice was deliberate: the study set out to assess what is possible when the entire stack, platform, model, and API, costs nothing. A commercial model such as GPT-4o would almost certainly perform better on harder reasoning tasks; however, selecting it would have bypassed the central research question. The trade-off was built into the evaluation design so that its consequences would be visible.

The conversation is organised into five thematic blocks: governance and board composition; workforce and gender; pay equity; ethnic diversity; and SDG disclosure. For each block, the agent collects raw figures from the user, computes the relevant percentages, and proceeds to the next block. Once all five blocks are complete, it presents a summary table for user confirmation and generates a structured report containing: an overall assessment assigning the firm to one of three performance clusters; a variable-by-variable comparison against sector benchmarks; an improvement plan with concrete targets and references to GRI 401-406, ESRS S1-S2, and SDGs 3, 5, and 10; a prioritised roadmap by time horizon; and a note on limitations. A batch mode is also available: if the user provides all data at once, the agent skips the dialogue and proceeds directly to the summary and report. Table 1 summarises the eleven variables processed by the agent, grouped by thematic block.

Table 1: Variables processed by the ESG advisory agent, grouped by thematic block.

Variable Group	Variables	Data Type	Benchmark Source
Board Composition	Board Gender Diversity, Executive Gender Diversity	Percentage (%)	Sector ESG cluster analysis
Workforce & Gender	Women Employees, Women Managers	Percentage (%)	Sector ESG cluster analysis
Pay Equity	Gender Pay Gap (index)	Index (0-100)	Provided pre-calculated index
Ethnic Diversity	Board, Executive & Workforce Ethnic Minorities	Percentage (%)	Sector ESG cluster analysis
Sustainability Disclosure	SDG Disclosure Score, Overall ESG Score	Score (0-100)	Sector ESG cluster analysis

3.3 Evaluation protocol

To evaluate the agent, a synthetic dataset of 40 fictitious energy companies was constructed. The firms are distributed across eight scenarios (five companies each), designed to stress-test different aspects of the agent's reasoning. Additional scenarios target specific edge cases, such as firms where board diversity is high but executive diversity is not, or vice versa. Each firm carries realistic values for all variables; crucially, the dataset includes both pre-calculated percentages and the raw figures from which those percentages were derived.

Two performance dimensions were measured. Value accuracy assesses whether the figure reported by the agent matches the dataset value within a tolerance of two percentage points. Judgment accuracy assesses whether the agent correctly identifies the direction of performance, that is, whether it accurately classifies a value as above or below the sector average.

4 Results

4.1 Overall accuracy

The results present a sharply divided picture. On eight of the eleven variables, Board Gender Diversity, Executive Gender Diversity, Women Employees, Women Managers, the three Ethnic Minorities indicators, and the Overall ESG Score, the agent achieves 98.1% value accuracy and 93.2% judgment accuracy. For a 17-billion-parameter model running on free infrastructure, these are strong results. The agent clearly understands what the variables mean, performs the required calculations correctly, and reliably compares results against benchmarks.

The remaining three variables tell a different story. Gender Pay Gap reaches only 33% value accuracy and 44% judgment accuracy. Board Cultural Diversity and Executive Cultural Diversity hover at approximately 42-45% on values. Something goes systematically wrong on these specific indicators; understanding why is

important both for improving the agent and for drawing broader conclusions about low-code AI systems. Table 2 presents the full accuracy results across all eleven variables.

Table 2: Value accuracy and judgment accuracy by variable.

Variable	Value Accuracy	Judgment Accuracy	Notes
Board Gender Diversity	98.1%	93.2%	Core variable - high performance
Executive Gender Diversity	98.1%	93.2%	Core variable - high performance
Women Employees	98.1%	93.2%	Core variable - high performance
Women Managers	98.1%	93.2%	Core variable - high performance
Ethnic Minorities (3 indicators)	98.1%	93.2%	Core variable - high performance
Overall ESG Score	98.1%	93.2%	Core variable - high performance
Gender Pay Gap	33%	44%	Conflict between raw data and index
Board Cultural Diversity	~42%	~44%	Confusion between similar variable populations
Executive Cultural Diversity	~45%	~44%	Confusion between similar variable populations

4.2 What went wrong, and why the model is responsible

The errors are not random. They follow two distinct patterns, both traceable to the same root cause: the model cannot reliably handle situations where the input contains competing signals that require prioritisation.

Consider the Gender Pay Gap. The test dataset provides the agent with two pieces of information: raw salary figures for women and men, and a pre-calculated index. The system prompt contains a general instruction, always compute percentages from raw numbers alongside a specific exception, for the Pay Gap, use the provided index. The model ignores the exception in every case and recalculates from raw salaries. Because the raw salaries and the index in the synthetic dataset were generated independently and do not perfectly correspond, the recalculated values diverge from the expected ones by as much as 78 percentage points.

Critically, this is not a prompt-writing problem. The instruction is present, clear, and unambiguous. The model simply cannot execute it. Recent work on instruction density by Jaroslawicz et al. (2025) explains why: the study identifies three degradation patterns as instruction complexity increases. Reasoning-oriented models maintain near-perfect adherence until reaching a threshold, then drop sharply. Mid-range models decline linearly. Llama 4 Scout specifically exhibits exponential decay, performance collapses rapidly once instruction complexity exceeds a modest level. The prompt used in this study, with its layered rules, benchmarks, formulas, and conversational flow structure, clearly crosses that threshold. Li et al. (2024), working on instruction-following robustness, reach a complementary conclusion: models with weaker contextual understanding tend to fixate on the most prominent instruction in the prompt rather than weighing competing directives against each other. In this case, the instruction to calculate from raw data is dominant; the exception for the Pay Gap is treated as noise.

The Cultural Diversity mix-ups fit the same explanation. Board Cultural Diversity and Executive Cultural Diversity use similar formulas but refer to different populations, board members versus executives respectively. Keeping them distinct requires the model to maintain two parallel variable tracks in working memory and route each raw figure to the correct denominator. This is precisely the kind of selective-attention task that mid-sized models struggle with when the prompt is already dense with other demands.

The 98.1% accuracy on core variables provides confidence that the architecture and prompt are not fundamentally at fault. The agent handles the same cognitive operations, percentage computation, benchmark comparison, qualitative labelling, almost flawlessly when the input is unambiguous. Failures are confined to cases requiring disambiguation or exception-handling: capacities that scale with model size.

4.3 Two attempts at fixing the Pay Gap

Two corrective strategies were tested on a subsample of six firms: three with large Pay Gap errors, two with moderate errors, and one control firm where the original output was already correct. Both strategies are consistent with the iterative logic of DSR.

Prompt-level fix: an explicit, prominent instruction was added: if both the index and the salary figures are present, always use the index and recalculate only when the index is missing. Result: one out of six firms was corrected. The model continued to recalculate from salaries in the remaining five cases. This is consistent with the Jaroslawicz et al. (2025) findings: adding instructions to an already overloaded prompt does not help a model past its instruction-following ceiling; if anything, it adds further complexity.

Input restructuring: instead of modifying the instruction, the source of conflict was removed. Raw salary figures were stripped from the input, leaving only the pre-calculated index. With no alternative figures available, the model had no opportunity to apply the wrong rule. Accuracy rose to 100% across all six firms. Table 3 summarises the two corrective attempts and their outcomes.

Table 3: Comparison of corrective strategies tested for the Gender Pay Gap variable.

Fix Attempt	Success Rate	Outcome	Interpretation
Prompt-level instruction added	1/6 firms	Largely failed	Model past its instruction-following ceiling; added instructions increase complexity without improving disambiguation
Input restructuring (raw salaries removed)	6/6 firms	100% accuracy	Eliminating the competing signal is more reliable than instructing the model to resolve it

The practical implication is clear. When working with a mid-sized open-source model, rewriting the prompt to resolve a conflict is often futile: the model lacks the reasoning depth to adjudicate between competing instructions regardless of how they are worded. Restructuring the input so that the conflict never arises is a more reliable strategy. For LLM-based tools intended to run on accessible, free infrastructure, the clarity of the input matters at least as much as the sophistication of the prompt.

5 Discussion and Conclusions

The results of this study provide several insights into the use of LLMs as tools for supporting knowledge creation from ESG data, particularly within low-code and accessible environments. The findings confirm that an ESG advisory agent built on a fully open and zero-cost infrastructure can produce structured and largely accurate outputs. The high value accuracy observed across most variables suggests that even mid-sized models are capable of performing basic analytical tasks, computing percentages, comparing values against sector benchmarks, with a high degree of reliability.

However, the results also highlight important limitations. The errors observed in specific variables, such as the Gender Pay Gap and Cultural Diversity indicators, reveal that the model struggles when the input contains competing signals or requires the application of exceptions to general rules. In these cases, the model tends to prioritise the most prominent instruction and ignore secondary ones, leading to systematic inaccuracies. This behaviour is consistent with recent studies on instruction-following limitations in LLMs, which show that model performance can degrade as instruction count and complexity increase (Jaroslawicz et al., 2025; Li et al., 2024).

From a theoretical perspective, these findings contribute to the literature on AI as a tool for knowledge creation and organisational sensemaking. While previous studies have emphasised the potential of AI to support interpretation and decision-making (Engstrom et al., 2024; Qin, 2024), this work shows that such potential is strongly dependent on the interaction between model capabilities, input structure, and system design. LLMs can effectively support the transformation of quantitative ESG data into qualitative insights when the analytical context is clear and unambiguous. Their contribution becomes less reliable when interpretation requires resolving conflicts, managing exceptions, or maintaining multiple parallel reasoning paths.

An important contribution of this study lies in the design implications emerging from the iterative DSR process. The comparison between the two corrective attempts for the Pay Gap issue shows that modifying the prompt is not always an effective strategy, especially when the model operates near its instruction-processing limits. Restructuring the input to eliminate ambiguity proved to be a more robust solution. This finding suggests a general design principle for low-code AI systems based on mid-sized models: reducing complexity at the input level is more effective than increasing complexity at the instruction level.

From a practical perspective, the study demonstrates that it is possible to develop accessible and configurable DSSs for ESG analysis using low-code tools and open-source models. This has important implications for organisations that lack the resources or technical expertise required to adopt more advanced AI solutions. The proposed approach allows users to combine quantitative data analysis with qualitative interpretation in a single workflow, supporting a more strategic use of ESG information. In this sense, the agent does not replace human decision-makers but acts as a support tool that enhances their ability to interpret data and identify potential improvement actions, in line with the traditional role of DSS (Poch et al., 2004).

The study acknowledges several limitations. First, the evaluation relies on a synthetic dataset which, although designed to be realistic, cannot fully capture the complexity of real-world organisational data. Second, the use of a mid-sized model (17 billion parameters) constrains the system's ability to handle complex reasoning tasks, particularly those involving ambiguity or conflicting information. Third, the quality of outputs remains

dependent on the quality and consistency of the input data, highlighting the importance of data preparation and validation.

Future research can build on these findings in several directions. First, the use of larger and more advanced models may improve performance in tasks requiring complex reasoning and exception handling. Second, testing the proposed system on real-world datasets would allow its applicability in organisational contexts to be assessed. Finally, further research could explore the extension of this approach to other ESG dimensions and the integration of additional data sources and analytical techniques.

In conclusion, this study shows that LLM-based agents, when integrated into low-code architectures, can play a valuable role in bridging the gap between quantitative ESG data and qualitative knowledge. While important limitations remain, particularly in relation to model capabilities and input design, the results highlight the potential of accessible AI solutions to support more informed and meaningful decision-making processes in the field of sustainability.

6 Ethics Declaration

The present research did not require the approval of an ethics committee, as it does not involve human subjects, sensitive personal data or experimentation on living beings. All the information used comes from publicly accessible sources and scientific literature.

7 AI Declaration

Generative Artificial Intelligence tools have not been used for the drafting or processing of the contents of this document. All sections of the paper were developed independently by the authors, without the aid of generative AI applications.

References

Bandi, A., Kongari, B., Naguru, R., Pasnoor, S. and Vilipala, S.V. (2025) “The rise of agentic AI: A review of definitions, frameworks, architectures, applications, evaluation metrics, and challenges”, Future Internet, Vol. 17, No. 9, 404.

Censi, R., Campana, P., Bellini, F., Schettino, F. and De Pucchio, C. (2025) “AI-Driven Process Mining for ESG Risk Assessment in Sustainable Management”, Buildings, Vol. 15, No. 23, 4260.

Censi, R., Campana, P., Schettino, F. and De Pucchio, C. (2025) “Corporate Sustainability Transition: Methodological Analysis for a Rating Model”, Corporate Social Responsibility and Environmental Management.

Cui, J. (2024) “Analysis of digital transformation on corporate ESG performance: A qualitative study”, Journal of Modern Social Sciences, Vol. 1, No. 2, pp. 89–98.

Delport, P.M., Von Solms, R. and Gerber, M. (2024) “Methodological Guidelines for Design Science Research”, Procedia Computer Science, Vol. 237, pp. 195–203.

Engstrom, A., Pittino, D., Mohlin, A., Johansson, A. and Edh Mirzaei, N. (2024) “Artificial intelligence and work transformations: integrating sensemaking and workplace learning perspectives”, Information Technology and People, Vol. 37, No. 7, pp. 2441–2461.

Gandia, J.A.G., de Lucas Ancillo, A. and del Val Nunez, M.T. (2025) “The Role of Artificial Intelligence and Knowledge in Enhancing Corporate Sustainability”, Journal of Innovation and Knowledge, Vol. 10, No. 5, 100792.

Govindan, K., Kaliyan, M., Kannan, D. and Haq, A.N. (2025) “Social sustainability”, Sustainable Supply Chain Management, Vol. 17, No. 8, pp. 455–466.

Graziani, M., Dutkiewicz, L., Calvaresi, D., Amorim, J.P., Yordanova, K. and Vered, M. et al. (2023) “A global taxonomy of interpretable AI: unifying the terminology for the technical and social sciences”, Artificial Intelligence Review, Vol. 56, No. 4, pp. 3473–3504.

Jaroslawicz, D. et al. (2025) How Many Instructions Can LLMs Follow at Once?, arXiv preprint.

Li, Z., Peng, B., He, P. and Yan, X. (2024) “Evaluating the Instruction-Following Robustness of Large Language Models to Prompt Injection”, Proceedings of EMNLP 2024, pp. 557–568.

Majekodunmi, T.G. (2025) “Embedding AI in ESG-Financial Reporting Systems to Enhance Non-Financial Disclosure Integrity, Assurance, and Investor Decision Relevance”, International Journal of Advance Research Publication and Reviews, Vol. 02, No. 6, pp. 464–484.

Mustafa, F., Smolarski, J. and Elamer, A.A. (2025) “The convergence of artificial intelligence and sustainability reporting: A systematic review of applications, challenges and future directions”, Business Strategy and the Environment, Vol. 34, No. 8, pp. 9761–9784.

Nielsen, C. (2023) “ESG reporting and metrics: From double materiality to key performance indicators”, Sustainability, Vol. 15, No. 24, 16844.

Poch, M., Comas, J., Rodriguez-Roda, I., Sanchez-Marre, M. and Cortes, U. (2004) “Designing and building real environmental decision support systems”, Environmental Modelling and Software, Vol. 19, No. 9, pp. 857–873.

Qin, X. (2024) “Intelligent technologies and methodological transformations in the social sciences”, Chinese Political Science Review, Vol. 9, No. 1, pp. 1–17.

Rahmani, A.M., Hemmati, A. and Abbasi, S. (2025) “The Rise of Large Language Models: Evolution, Applications, and Future Directions”, Engineering Reports, Vol. 7, No. 9, e70368.

Sierdovski, M., Pilatti, L.A. and Rubbo, P. (2022) “Organizational competencies in the development of environmental, social, and governance (ESG) criteria in the industrial sector”, Sustainability, Vol. 14, No. 20, 13463.

Taulli, T. and Deshmukh, G. (2025) Building Generative AI Agents: Using LangGraph, AutoGen, and CrewAI, Apress.

Verschuren, P. and Hartog, R. (2005) “Evaluation in design-oriented research”, Quality and Quantity, Vol. 39, No. 6, pp. 733–762.

Zong, Z. and Guan, Y. (2025) “AI-driven intelligent data analytics and predictive analysis in Industry 4.0: Transforming knowledge, innovation, and efficiency”, Journal of the Knowledge Economy, Vol. 16, No. 1, pp. 864–903.

Can Ethics Become Isomorphic? Informational Power in GenAI Governance

Marco Di Giovanni
University G. d'Annunzio, Pescara, Italy

Abstract

The rapid diffusion of Generative AI (GenAI) across organizational contexts is transforming not only knowledge work and creative practices, but also the way ethical questions are defined, structured, and legitimized at the field level. Despite the growing scholarly attention devoted to AI ethics, relatively little is known about how the moral boundaries of GenAI are stabilized and diffused across dominant providers. This paper argues that the publicly available governance artifacts produced by leading GenAI providers — including usage policies, safety frameworks, transparency statements, and model documentation — can be understood as moral ordering infrastructures that shape what is permitted, what is prohibited, which harms are foregrounded, and how responsibility is framed.

Drawing on Information Ethics and the concept of the infosphere, the paper proposes that those who govern the infrastructures of generative intelligence exercise a form of informational power capable of orienting the moral grammar through which organizations interpret and manage ethical questions. Integrating this perspective with neo-institutional theory, the paper introduces the concept of ethical isomorphism: a process through which coercive, mimetic, and normative pressures, combined with the concentration of informational power, foster convergence in ethical governance scripts across dominant providers.

The paper develops this conceptual argument through a focused comparative reading of ten publicly available governance documents produced by five leading providers — OpenAI, Anthropic, Google, Microsoft, and Meta. The comparison uses a structured analytical template covering ethical principles, harm taxonomies, stakeholder framing, trade-offs, and enforcement and accountability mechanisms. This approach allows the paper to examine how abstract AI ethics principles are translated into observable governance scripts across the GenAI provider field.

The comparative analysis reveals a pattern of partial but meaningful convergence. The strongest alignment emerges in harm taxonomies, prohibited-use structures, and governance architectures, while the main divergences concern governance style, the breadth of stakeholder framing, and the degree to which moral trade-offs are explicitly acknowledged. The paper concludes that the public governance artifacts of leading providers make visible an emerging form of ethical isomorphism — not as full normative identity, but as a progressive standardization of the way ethics is operationalized and institutionalized. Such convergence may narrow the ethical pluralism available to adopting organizations, which may find themselves importing preconfigured moral assumptions alongside GenAI tools. The paper offers relevant implications for research, organizations, and AI governance policy.

Keywords: Generative AI governance; ethical isomorphism; informational power; algorithmic governance; ethical pluralism

Paper type: Academic Research Paper

1 Introduction

The rapid diffusion of Generative AI (GenAI) is reshaping organizational knowledge practices, creative work, and decision-making processes, while simultaneously intensifying ethical concerns related to bias, opacity, privacy, misuse, and accountability. Recent studies show that GenAI is increasingly embedded in knowledge-intensive work and human–AI collaboration, making it far more than a peripheral technological tool (Anthony et al., 2023; Benbya et al., 2024; Brynjolfsson et al., 2025). Yet the problem is not only that GenAI creates new ethical risks. It is also that the infrastructures through which these systems are governed are becoming increasingly central to how organizations define what counts as acceptable use, relevant harm, and legitimate control. In digitally mediated organizational environments, information is not merely an input to action, but a constitutive element of socio-economic life and inter-organizational relations (Vaccaro & Madsen, 2009). This makes the governance of GenAI inseparable from the governance of the informational environment in which organizations operate.

Despite the rapid expansion of the AI ethics literature, much of the debate has focused on the proliferation of principles and on the gap between ethical aspiration and practical implementation (Jobin et al., 2019; Morley et al., 2021). What remains less understood is how the ethical boundaries of GenAI are stabilized at field level by the actors who control its most influential infrastructures. In practice, these boundaries are increasingly articulated through provider-produced governance artifacts—such as usage policies, safety frameworks,

transparency statements, and model documentation—that specify what is allowed, what is prohibited, which harms matter, and how responsibility is framed. These artifacts can be understood as institutional devices that help structure the moral and epistemic boundaries of the field, much like ethical codes can operate as scripts that standardize organizational understandings of legitimacy and responsibility (Novis-Deutsch et al., 2022).

This paper approaches that issue through the lens of Information Ethics and the concept of the infosphere. The infosphere can be understood as the broader informational environment within which human and technological agents produce, exchange, and transform information (Floridi, 2009, 2010, 2013). In this view, organizations are informational actors embedded in a socio-digital environment whose structure shapes knowledge, judgment, and governance. This perspective expands the ethical focus from the use of technology to the governance of informational infrastructures. When information becomes constitutive of organizational reality, governing informational infrastructures also means shaping the conditions under which harms are classified, values are prioritized, and moral judgments become legitimate. In this sense, informational power is not only technical or economic; it is also moral and epistemic. GenAI is especially significant because it does not merely operate within the infosphere, but actively contributes to organizing it. Under conditions of provider concentration, those who govern generative infrastructures may also shape the moral grammar through which organizations interpret and manage ethical questions.

This concern becomes sharper when viewed through a neo-institutional lens. Organizational fields do not converge only because actors share technologies; they also converge through coercive, mimetic, and normative pressures that stabilize legitimate vocabularies, categories, and governance practices (DiMaggio & Powell, 1983; Boxenbaum & Jonsson, 2008). In the GenAI provider field, this raises a more specific question: do public governance artifacts progressively standardize moral categories, stakeholder framings, enforcement logics, and acceptable trade-offs? This paper refers to that possibility as ethical isomorphism: a process through which concentrated informational power and institutional pressures foster convergence in ethical governance scripts across dominant providers. Such convergence matters because it may reduce ethical pluralism, that is, the diversity of moral perspectives through which organizations and stakeholders can engage in reflexive and legitimate ethical deliberation (Buchholz & Rosenthal, 1996; Greenwood & Freeman, 2011). If ethics becomes increasingly procedural, standardized, and weakly contestable, organizations may end up importing preconfigured moral assumptions together with GenAI tools.

The paper develops this argument through three related conceptual and exploratory questions: (RQ1) How can ethical isomorphism be theorized at the intersection of Information Ethics and neo-institutional theory? (RQ2) How are ethical governance scripts articulated across the public governance artifacts of leading GenAI providers? (RQ3) How does convergence in these scripts reshape the space for ethical pluralism in organizational GenAI adoption? To address these questions, the paper combines conceptual theory-building with a focused comparative reading of public governance artifacts produced by leading GenAI providers. By comparing how providers define principles, prioritize harms, frame stakeholders, and construct enforcement and accountability, the paper examines how abstract AI ethics principles are translated into observable governance scripts. In doing so, it contributes to the literature in three ways. First, it reframes GenAI ethics as a field-level institutional outcome rather than merely a set of abstract principles. Second, it integrates Information Ethics and neo-institutional theory to explain how informational power may translate into convergent ethical governance. Third, it highlights the practical and policy relevance of preserving ethical pluralism in an increasingly concentrated GenAI landscape.

2 Theoretical Background

2.1 Information ethics and the infosphere

To understand why GenAI governance matters ethically, it is useful to begin from the perspective of Information Ethics. Rather than treating information as a neutral input or as a mere technical resource, Information Ethics conceives it as a constitutive element of contemporary socio-digital reality (Floridi, 2009, 2010, 2013). This broader informational environment can be described as the infosphere: the relational space in which human and technological agents continuously generate, exchange, process, and transform information. In this view, organizations are not external observers of the digital world; they are informational actors embedded in it, whose practices, decisions, and governance arrangements are increasingly shaped by informational infrastructures. Ethical questions, therefore, do not arise only at the level of individual conduct or technological malfunction, but at the level of how informational environments are structured, maintained, and legitimized.

This perspective implies that ethics is not simply a matter of applying abstract normative principles after technological systems have been developed. Rather, ethics becomes a structural issue concerning the integrity, quality, and sustainability of informational relations. Preserving the integrity of the infosphere means preserving the conditions under which knowledge can be produced, shared, and contested without being excessively distorted, monopolized, or procedurally narrowed (Floridi, 2009). From this standpoint, the governance of information is also a governance of moral possibility: actors who shape informational conditions indirectly shape what becomes visible, legitimate, and actionable in organizational life. This is precisely why the rise of GenAI deserves attention not only as an innovation in productivity or creativity, but also as a transformation in the moral architecture of the informational environment itself.

2.2 Informational power and the rise of generative AI

The relevance of this perspective has become especially clear with the rapid diffusion of Generative AI across organizational contexts. GenAI systems are now increasingly involved in creative work, decision support, communication, and knowledge-intensive practices (Anthony et al., 2023; Benbya et al., 2024; Brynjolfsson et al., 2025). Their significance, however, exceeds mere adoption rates. GenAI systems are trained on vast informational corpora, operate by processing and recombining information at scale, and continuously influence how information is retrieved, framed, and redistributed. In this sense, GenAI does not simply function within the infosphere; it contributes to organizing it. The more such systems mediate everyday organizational knowledge work, the more they participate in shaping the epistemic and moral conditions under which actors interpret problems, define harms, and judge acceptable conduct.

This leads to the notion of informational power. Following Floridi (2009, 2010, 2013), informational power can be understood as the capacity to shape moral and epistemic boundaries through control over informational flows and over the conditions of their production and circulation. Such power is not merely technical. It is also epistemic, because it influences what counts as relevant or credible knowledge, and moral, because it affects the classification of risks, acceptable uses, prohibited practices, and legitimate trade-offs. In the GenAI field, this power is concentrated in a relatively small set of dominant providers whose models, interfaces, safety policies, and usage rules increasingly function as infrastructural conditions for adoption. As a result, provider governance does not only regulate products; it may also shape the broader “space of the thinkable” available to organizations using those products.

2.3 Institutional mechanisms, ethical isomorphism, and ethical pluralism

While Information Ethics helps explain why control over informational infrastructures matters, neo-institutional theory helps explain how convergence may emerge across actors operating in the same field. Classic institutional theory shows that organizations tend to become similar through coercive, mimetic, and normative pressures (DiMaggio & Powell, 1983; Boxenbaum & Jonsson, 2008). Coercive pressures stem from formal and informal obligations, including regulatory expectations and compliance demands. Mimetic pressures arise under uncertainty, when actors imitate highly legitimate models. Normative pressures emerge through professional standards, expert communities, auditing practices, and shared vocabularies of appropriateness. In highly visible and contested technological fields, these pressures may not only shape organizational structures, but also the categories through which ethical legitimacy is articulated and defended.

Applied to GenAI governance, this suggests that providers may progressively align not only on technical safeguards but also on the ethical language embedded in their public artifacts. Governance documents can standardize what counts as harm, which stakeholders are rendered visible, how accountability is framed, and which ethical trade-offs are treated as legitimate or unavoidable. This paper refers to such a process as ethical isomorphism: the alignment of moral logics and governance scripts across actors within the infosphere under conditions of concentrated informational power and institutional pressure. Importantly, the object of convergence here is not formal organizational structure alone, but the institutionalization of “ethical alignment” itself as a field-level governance script (Wijen, 2014; Islam & Greenwood, 2024). In other words, ethics may become standardized not because moral dilemmas have been resolved, but because certain templates of responsible conduct become widely diffused and difficult to contest.

This theoretical possibility foregrounds a final construct: ethical pluralism. Ethical pluralism refers to the coexistence of multiple moral perspectives that enable reflexive deliberation, stakeholder responsiveness, and legitimate organizational judgment (Buchholz & Rosenthal, 1996; Greenwood & Freeman, 2011). In heterogeneous organizational environments, pluralism is not a weakness but a necessary condition for addressing complex ethical trade-offs. If dominant GenAI providers converge on a narrow moral grammar—particularly one that becomes increasingly procedural, compliance-oriented, and weakly reflexive—then ethical

pluralism may erode even as ethics discourse expands. The problem is therefore not simply whether providers use ethics language, but whether the concentration of informational power contributes to reducing the diversity of moral frameworks through which organizations can interpret and govern GenAI. This is the theoretical basis on which the present study investigates public governance artifacts of leading providers.

3 Method

This paper adopts a conceptual theory-building approach supported by a focused comparative analysis of public governance artifacts produced by leading Generative AI providers. The aim is to develop the concept of ethical isomorphism and to examine how providers define, justify, and operationalize the ethical boundaries of their systems in publicly available documents. Ethics is treated here as an institutionalized form of governance, expressed through principles, harms, stakeholder framings, trade-offs, and accountability mechanisms.

The empirical corpus consists of ten publicly available governance documents produced by five dominant providers: OpenAI, Anthropic, Google, Microsoft, and Meta. The sample was selected on the basis of field relevance, organizational visibility, and the availability of comparable public governance documentation. The corpus includes the following documents: for OpenAI, *Usage Policies* and *Model Spec*; for Anthropic, *Transparency Hub* and *Responsible Scaling Policy Version 3.0*; for Google, *Gemini for Google Cloud and Responsible AI* and the *Generative AI Prohibited Use Policy*; for Microsoft, *Responsible AI: Ethical policies and practices* and the *Transparency Note for Microsoft Copilot*; and for Meta, the *Responsible Use Guide* and the *Llama 4 Acceptable Use Policy*.

Each document was treated as a governance artifact, and each provider as a comparative case. The analysis followed a structured analytical template covering: document type, stated purpose, explicit ethical principles, prioritized harms, marginal or weakly articulated harms, stakeholder visibility, trade-offs, enforcement and accountability mechanisms, degree of operationalization, compliance language, references to standards or evaluation practices, and indicators of ethical openness or closure. Cross-case comparison focused on five dimensions: (1) ethical principles and their operationalization; (2) harm taxonomies; (3) stakeholder framing; (4) trade-offs between safety, openness, access, and responsibility; and (5) enforcement and accountability. This comparative structure supports the distinction between rhetorical similarity and more substantive alignment across providers.

The analytical process combined author-led document reading with AI-assisted comparative synthesis. The documents were read and interpreted by the author, while ChatGPT was used to support the identification of cross-case patterns of convergence, divergence, and ethical pluralism. The AI-assisted outputs were used as interpretive inputs within the author's analytical process and were reviewed, refined, and integrated through direct engagement with the original documents.

The analysis is theory-driven. Recurring similarities across provider artifacts are interpreted through neo-institutional theory as indicators of coercive, mimetic, and normative pressures within a concentrated provider field. Divergences across providers are analyzed as important variations in governance style, stakeholder framing, and ethical openness. This approach allows the paper to assess how public governance discourse makes ethical isomorphism visible as an emerging pattern in the GenAI provider field, and how such convergence may reshape ethical pluralism in the infosphere.

4 Findings

The comparative analysis reveals a visible pattern of partial convergence across the public governance artifacts of leading GenAI providers. Across the corpus, ethics is consistently framed through a shared vocabulary centered on safety, transparency, accountability, privacy, and responsible use. This convergence is visible not only at the level of declared principles, but also in the way providers classify harms, define prohibited uses, and operationalize governance through safeguards, monitoring, and restrictions. This convergence remains partial: providers differ in governance style, in the explicit treatment of trade-offs, and in the breadth of stakeholder framing. Table 1 summarizes the principal areas of cross-provider convergence and divergence emerging from the comparative analysis.

Table 1: Main cross-provider convergences and divergences in ethical governance

Dimension	Main convergence	Main divergence	Analytical implication
Ethical principles	All providers invoke safety, transparency, accountability, privacy, and responsible use.	OpenAI emphasizes model behavior; Anthropic emphasizes frontier safety; Microsoft emphasizes institutional governance; Meta emphasizes developer responsibility and open innovation; Google emphasizes prohibited use and risk control.	Strong rhetorical convergence with differentiated governance styles.
Harm taxonomy	Broad overlap on illegal use, child safety, self-harm, privacy misuse, cyber misuse, and harmful content.	Anthropic gives more weight to catastrophic risks; Google is especially explicit on restricted high-impact uses.	Strongest area of substantive convergence.
Trade-offs	Utility and risk control are salient across all cases.	OpenAI and Meta articulate trade-offs more explicitly; Microsoft, Google, and Anthropic frame them more procedurally.	Ethical pluralism is present, but unevenly articulated.
Enforcement and accountability	All providers rely on some mix of restrictions, safeguards, filtering, evaluation, or monitoring.	Enforcement architectures differ: behavioral hierarchy (OpenAI), scaling safeguards (Anthropic), whole-system governance (Microsoft), developer-layered controls (Meta), prohibited-use closure (Google).	Strong procedural convergence despite organizational variation.
Stakeholder framing	Users and developers are visible in all cases.	Microsoft offers the broadest stakeholder view; Meta gives more discretion to developers; participatory stakeholder voice remains weak across the corpus.	Managed stakeholder visibility is common; strong stakeholder agency is rare.
Ethical pluralism	Ethics is present across the field.	Most documents proceduralize ethics into safety and compliance rather than opening sustained space for contestation.	The field shows partial ethical closure rather than rich pluralism.

4.1 Convergence in ethical principles and harm framings

The strongest cross-case convergence appears in the articulation of ethical principles and harm categories. All five providers frame ethics through a common responsible-AI vocabulary and repeatedly prioritize a similar core of harms: illegal use, child sexual exploitation, self-harm, privacy and data misuse, cybersecurity abuse, and harmful or unsafe content. Google's prohibited-use policy is especially explicit in codifying restrictions, particularly in high-impact contexts. Microsoft's transparency note operationalizes risks through concrete mitigation logic. Meta's responsible-use materials connect harms to layered safeguards and developer controls. OpenAI combines policy prohibitions with model behavior rules, while Anthropic extends the harm horizon toward frontier and catastrophic risk. The analytical implication is that providers are not only using similar language, but are stabilizing a comparable moral core of what must be prevented, restricted, or controlled.

4.2 Convergence in governance architecture

A second clear pattern is that ethics is translated into governance architecture rather than left at the level of principle statements. Across the corpus, providers operationalize ethics through recurring devices: prohibited-use lists, behavior specifications, evaluations, safeguards, monitoring, reporting channels, and access restrictions. Here, convergence becomes operational. Microsoft uses a whole-system governance logic with product-level transparency and risk controls. Anthropic links capability development to escalating safeguards and transparency commitments. Meta adopts a layered mitigation approach that spans model-level and system-level controls. OpenAI structures governance through a behavioral hierarchy and refusal logic. Google relies most visibly on prohibition and human-supervision requirements. These architectures differ in form, but they perform a similar function: they convert ethics into actionable control.

4.3 Divergence in trade-offs, stakeholder framing, and pluralism

The main divergences concern how providers frame stakeholders and how explicitly they acknowledge moral trade-offs. Microsoft offers the broadest stakeholder formulation by including affected people and deployment context. Meta assigns greater responsibility to developers and downstream builders. OpenAI centers users and developers, while Anthropic gives more visibility to expert communities and society-wide risk. Google's materials frame stakeholders mainly through protection and oversight. Trade-offs are also unevenly articulated: OpenAI and Meta are the clearest in acknowledging tensions between usefulness and safety, or alignment and

helpfulness, whereas Anthropic, Microsoft, and Google frame these tensions more procedurally through risk management and compliance. The overall result is a condition of limited ethical pluralism. Ethics is visible across the corpus, but it is largely expressed through a standardized governance grammar centered on safety, mitigation, and control, with less space for explicit moral contestation or competing ethical perspectives.

5 Discussion

The comparative analysis substantiates the core argument of the paper: the ethics of GenAI is increasingly institutionalized through provider-produced governance artifacts that define the moral and operational boundaries of use. Across the five providers, ethics is translated into a relatively stable governance grammar organized around safety, transparency, accountability, privacy, harm prevention, and controlled use. Provider documents therefore operate as infrastructures of ethical ordering within the infosphere, shaping what becomes permissible, risky, contestable, or prohibited in organizational use of GenAI.

From a neo-institutional perspective, the pattern observed can be interpreted as partial ethical isomorphism. The providers do not present identical moral frameworks; rather, they converge on a common procedural architecture of governance. The strongest alignment appears in harm taxonomies, prohibited-use structures, and enforcement logics, while divergence remains more visible in governance style, stakeholder breadth, and the explicit treatment of trade-offs. This means that convergence in the field is primarily a matter of increasingly similar ways of operationalizing ethics through recurring categories, safeguards, and control mechanisms. In that sense, the field is becoming more alike because providers are stabilizing a common model of what responsible GenAI governance should look like.

This result is especially important when interpreted through the lens of Information Ethics. As GenAI providers increasingly mediate how information is generated, filtered, and distributed, their governance scripts also shape the moral boundaries of the infosphere. The concentration of informational power in a limited number of providers therefore has ethical significance beyond product design or compliance. It affects which harms are foregrounded, which stakeholders are rendered visible, which trade-offs are normalized, and which forms of conduct are treated as legitimate or unacceptable. The analysis suggests that this power operates through standardized governance templates that consolidate ethics into a portable grammar of responsible AI.

The findings also clarify the paper's position on ethical pluralism. Ethics is highly visible across the corpus, but its dominant form is procedural, safety-centered, and compliance-compatible. Stakeholders are usually framed as subjects of protection rather than as actors with strong voice in governance. Trade-offs are acknowledged unevenly, and explicit moral contestation remains limited. The current governance landscape therefore reflects a condition of compressed pluralism: variation remains, but it is increasingly contained within a shared procedural grammar that may narrow the range of ethical frameworks available to organizational adopters.

For organizations, this has a direct implication. Adopting GenAI means adopting technical systems together with provider-defined governance scripts. These scripts can increase safety and standardization, while also shaping the ethical categories through which organizations interpret responsibility, risk, and acceptable use. For researchers and policymakers, the implication is equally clear: the analysis of GenAI ethics should move beyond lists of principles and focus more closely on how ethical governance is concretely structured, diffused, and normalized across the provider field.

6 Conclusion

This paper has shown that the ethics of GenAI is increasingly articulated through provider-produced governance artifacts that define what counts as acceptable use, relevant harm, legitimate control, and responsible deployment. Across the five providers, the analysis identifies a condition of partial but meaningful convergence. The strongest alignment appears in ethical vocabulary, harm taxonomies, and governance mechanisms, while divergence remains more visible in governance style, stakeholder breadth, and the explicit treatment of trade-offs. The main conclusion is therefore that the public governance artifacts of leading providers make visible an emerging form of ethical isomorphism: not full normative identity, but a growing convergence in the way ethics is operationalized and institutionalized across dominant GenAI providers.

The contribution of the paper is twofold. First, it shifts the analysis of GenAI ethics from abstract principles to public governance scripts, showing that ethics is structured through operational categories, restrictions, safeguards, and accountability devices. Second, it connects Information Ethics and neo-institutional theory to show how concentrated informational power can help stabilize a shared procedural moral grammar in the infosphere. In this sense, the paper suggests that the central issue is no longer only whether providers adopt

ethical principles, but how a small number of influential actors increasingly shape the boundaries of what becomes ethically thinkable and governable for organizations adopting GenAI.

The study also opens a research agenda on the relationship between public governance scripts, organizational adoption, and actual enforcement practices. Future research can extend the analysis through interviews, internal governance evidence, external audits, and enforcement cases in order to examine how public ethical scripts are translated into organizational practice and whether the convergence observed here deepens over time. Preserving ethical pluralism in GenAI governance is therefore a central challenge for organizations, policymakers, and researchers concerned with the future of knowledge work in an increasingly provider-mediated infosphere.

7 Ethics declaration

This study is based exclusively on the analysis of publicly available documents produced by GenAI providers. It does not involve human participants, personal data, interviews, experiments, or other procedures requiring formal ethical approval. Accordingly, no institutional ethics review was required.

8 AI declaration

AI tools were used during manuscript preparation to support document comparison, the coding of cross-case similarities and differences, and the refinement of language. In particular, ChatGPT was used to assist the analysis of convergence, divergence, and ethical pluralism across provider governance documents. All AI-assisted outputs were reviewed, checked against the original sources, and validated by the author, who takes full responsibility for the final analysis, interpretations, and wording of the paper.

References

Anthony, C., Bechky, B.A. and Fayard, A.-L. (2023) "Collaborating" with AI: Taking a system view to explore the future of work, *Organization Science*, Vol. 34, No. 5, pp 1672–1694.

Benbya, H., Strich, F. and Tamm, T. (2024) "Navigating generative artificial intelligence: Promises and perils for knowledge and creative work", *Journal of the Association for Information Systems*, Vol. 25, No. 1, pp 23–36.

Boxenbaum, E. and Jonsson, S. (2008) "*Isomorphism, diffusion and decoupling*", in Greenwood, R., Oliver, C., Suddaby, R. and Sahlin, K. (Eds.), The SAGE Handbook of Organizational Institutionalism, SAGE Publications, London, pp 78–98.

Brynjolfsson, E., Li, D. and Raymond, L. (2025) "Generative AI at work", *Quarterly Journal of Economics*, Vol. 140, No. 2, pp 889–942.

Buchholz, R.A. and Rosenthal, S.B. (1996) "Toward a New Understanding of Moral Pluralism", *Business Ethics Quarterly*, Vol. 6, No. 3, pp 263–275.

DiMaggio, P.J. and Powell, W.W. (1983) "The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields", *American Sociological Review*, Vol. 48, No. 2, pp 147–160.

Floridi, L. (2009) "Network Ethics: Information and Business Ethics in a Networked Society", *Journal of Business Ethics*, Vol. 90, No. 4, pp 649–659.

Floridi, L. (2010) *Information: A Very Short Introduction*, Oxford University Press, Oxford.

Floridi, L. (2013) *The Ethics of Information*, Oxford University Press, Oxford.

Greenwood, M. and Freeman, R.E. (2011) "Ethics and HRM: The Contribution of Stakeholder Theory", *Business and Professional Ethics Journal*, Vol. 30, No. 3/4, pp 269–292.

Islam, G. and Greenwood, M. (2024) "Generative artificial intelligence as hypercommons: Ethics of authorship and ownership", *Journal of Business Ethics*, Vol. 192, No. 3, pp 659–663.

Jobin, A., Ienca, M. and Vayena, E. (2019) "The global landscape of AI ethics guidelines", *Nature Machine Intelligence*, Vol. 1, No. 9, pp 389–399.

Morley, J., Floridi, L., Kinsey, L. and Elhalal, A. (2021) "From what to how: An initial review of publicly available AI ethics tools, methods and research to translate principles into practices", *Science and Engineering Ethics*, Vol. 27, No. 1, pp 1–31.

Novis-Deutsch, N., Rom, S. and Ben-Ari, R. (2022) "A shared language? Ethical codes as institutional scripts in organizational ethics", *Journal of Business Ethics*, Vol. 178, No. 3, pp 627–642.

Vaccaro, A. and Madsen, P. (2009) "Corporate dynamic transparency: The new ICT-driven ethics?", *Ethics and Information Technology*, Vol. 11, No. 2, pp 113–122.

Wijen, F. (2014) "Means versus ends in opaque institutional fields: Trading off compliance and achievement in sustainability standards", *Academy of Management Review*, Vol. 39, No. 3, pp 302–323.

A Regenerative Framework for Socio-Ecological Resilience: The Case of Armani's Apulia Regenerative Cotton Project

Marta Menegoli[1], Luigi Jesus Basile[2], Roberto Mauriello[3]
[1] University of Salento, Lecce, Italy
[2] LITEM, Institut Mines-Telecom Business School, Paris, France
[3] Pegaso University, Naples, Italy

Abstract

In recent years, increasing stress on socio-ecological systems has highlighted the limits of traditional sustainability approaches, fostering growing interest in regenerative perspectives aimed at restoring natural assets and supporting long-term resilience. These approaches may be particularly relevant for the fashion industry, given the environmental pressures associated with production activities and the complexity of global supply chains. However, regenerative approaches are still nascent in the fashion industry, and the literature offers limited insight into how such models emerge and consolidate over time. To address this gap, this study analyses the development of a regenerative initiative in the fashion industry through the lens of Institutional Theory. This perspective is particularly suitable to examine how coercive, normative, and mimetic pressures affect the different actors involved and how these pressures are interpreted and translated into practices over time, contributing to the institutionalization of new business models based on regenerative approaches.

From a methodological perspective, the study builds on a longitudinal single embedded case study of the Apulian Regenerative Cotton Project promoted by Armani, a leading luxury fashion brand. The analysis considers the key actors involved in the initiative, including the focal firm, scientific partners, and local agricultural firms across the supply chain, following the project from its inception to garment production. The findings suggest that normative pressures play a central role in orienting actors' behaviours and in defining appropriate practices throughout the initiative. Mimetic pressures emerge as the initiative gains visibility, while coercive pressures remain relatively limited, although potentially evolving in connection with emerging regulatory frameworks. Overall, the institutionalization process appears as gradual and non-linear, shaped by the continuous interpretation and translation of multiple pressures into concrete practices through the interaction among actors. Overall, this study contributes to the literature by providing an empirical analysis of a regenerative initiative in the fashion industry through the lens of Institutional Theory, showing how regenerative business models may emerge and diffuse. From a practical perspective, the findings highlight the importance of recognising and managing multiple institutional pressures, as well as supporting coordination among heterogeneous actors involved in the development of such initiatives.

Keywords: Regenerative approaches; Socio-ecological systems; Institutional Theory; Fashion industry; Regenerative business models

Paper type: Academic Research Paper

1 Introduction

In recent years, the urgency of addressing systemic stress in socio-ecological systems has become increasingly evident (Gualandris et al., 2024). Climate change and the degradation of natural resources increased the exposure of socio-ecological systems to risks and disruptions, while also revealing structural vulnerabilities that require moving beyond traditional approaches to sustainability (Erzurumlu et al., 2023). In this context, regenerative approaches, grounded in the need to restore natural assets, protect diversity, and foster distributed forms of governance, are attracting growing attention from both policymakers and scholars (Gugganig, 2025). In the EU, recent regulatory developments started to explicitly recognize regenerative practices. The European Regulation 2024/3012 introduces the Carbon Removals Certification Framework (CRCF), establishing a formal basis for certifying carbon removal activities. The delegated acts adopted in 2026 further specify technical standards for carbon farming, identifying core practices, and measures for certification. On a broader perspective, the EU Nature Restoration Law, as well as consolidated frameworks like the Task Force on Nature-related Financial Disclosures (TNFD) and the Global Reporting Initiative (GRI), while primarily focused on biodiversity and sustainability reporting, highlight the central role of restoration interventions, specifying relevant domains, metrics, and interventions.

On this other hand, recent studies examine how regenerative agriculture and restoration approaches are articulated into sets of practices aimed at improving soil health, enhancing biodiversity, and supporting the long-term resilience of socio-ecological systems (Beacham et al., 2023). Thus, diversification, reduced input intensity, and soil regeneration are increasingly discussed as key components of more sustainable and adaptive agricultural systems (Shennan-Farpón et al., 2025). These approaches are also associated with the

reconfiguration of production systems that support sustainability through adaptive and resilient practices, and are increasingly framed as part of broader transitions toward more system-oriented models of agricultural production (Shahmohamadloo et al., 2021).

These approaches may be particularly relevant for the fashion industry, given the environmental stress that production activities impose on local territories, and the delicate social issues affecting global fashion supply chains (Colasante et al., 2025; Peter John and Mishra, 2023). In this context, regenerative approaches may be instrumental in enabling the shift towards more sustainable and responsible business practices.

Despite this, the application of regenerative approaches in the fashion industry remains limited and the literature is still in early stages. Available studies, however, suggest the potential of sustainability-oriented sourcing practices and production systems for the fashion industry, particularly in relation to natural fibres and circular production processes (Brydges, 2021; Oliveira Duarte et al., 2022). Such initiatives involve multiple actors, including farmers, manufacturers, and supporting institutions, and require strong coordination across the supply chain to develop shared business models and align diverse objectives.

In this scenario, this study aims to contribute by providing a longitudinal analysis of one of the most recent and relevant applications of regenerative agricultural approaches in the fashion industry, namely the Apulian Regenerative Cotton Project, coordinated by Armani, a leading luxury fashion brand. To do so, the study adopts Institutional Theory as a theoretical lens. This is particularly suitable to examine coercive, normative, and mimetic pressures to which the different actors involved in the initiative are exposed across the different phases of the project (DiMaggio and Powell, 1983). From this perspective, the Institutional Theory can help explain how each actor interprets and responds to coercive pressures, arising from formal rules and regulatory requirements, normative pressures, related to shared norms and expectations, and mimetic pressures, to achieve and maintain legitimacy at each stage of the project. Moreover, this perspective allows us to understand how actors contribute to the institutionalization of practices and novel regenerative business models in a continuously evolving context (Zucker, 1987). In doing so, it provides insights into how such practices can be progressively established and diffuse, becoming a reference point for similar initiatives within the industry. Consistently, the study addresses the following research questions:

- *RQ1) What types of institutional pressures affect the different actors involved in regenerative initiatives within the fashion industry?*
- *RQ2) How do these actors interpret and respond to such pressures across the different phases of the initiative?*

To address these questions, the study adopts a qualitative single case study design. The Apulian Regenerative Cotton Project represents a particularly suitable case, as one of the first and most relevant applications of regenerative approaches in the fashion industry. The study follows an embedded design and thus considers multiple units of analysis, including the leading firm, Armani, as well as the other key actors involved in the initiative namely scientific partners, local farmers, suppliers, and certification bodies. This allows to capture the different pressures experienced and enacted by the various actors across the supply chain. Moreover, the study adopts a longitudinal perspective, analysing the initiative across its key phases, to investigate how pressures develop over time and how practices are institutionalised by the partners to carry forward the project. Data collection is based on secondary sources and semi-structured interviews conducted with key actors involved in the project, covering the different stages of the initiative.

This study contributes to the literature by analysing an innovative, regenerative project in the fashion industry through the lens of Institutional Theory. In doing this, the study helps to explain how institutional pressures, including coercive, normative, and mimetic, affect key actors across the supply chain and shape their behaviour throughout the different phases of the initiative. Furthermore, the study shows how these pressures are interpreted and translated into practices that are progressively institutionalized over time, thus showing how regenerative business models can emerge and diffuse within the fashion industry, offering a grounded reference for similar initiatives. From a practical perspective, the findings highlight the importance of recognising and managing the multiple pressures affecting different actors, as well as the need to support coordination and align diverse objectives. In this sense, the study provides actionable insights for practitioners involved in the design and implementation of regenerative initiatives, supporting the development and scaling of such models within the industry.

The remainder of the paper is structured as follows. Section 2 establishes the theoretical framework, integrating Institutional Theory with the concept of socio-ecological regeneration to examine how isomorphic pressures drive organizational legitimacy. Section 3 details the research methodology, outlining the exploratory qualitative approach and the longitudinal single embedded case study design focused on the Apulian Regenerative Cotton Project. Section 4 presents the results and discussions, tracing the initiative through four key phases, initiation, experimentation, structuring, and diffusion, and mapping the evolution of institutional

pressures throughout this process. Section 5 explores the theoretical and practical implications, offering a narrative on institutional entrepreneurship and a blueprint for supply chain coordination. Finally, Section 6 provides the conclusions, summarizes the study's contributions, and addresses its limitations while suggesting directions for future research.

2 Theoretical Framework: Institutional Theory and Regeneration

To understand how regenerative practices gain legitimacy and diffuse across industries, this study employs Institutional Theory (DiMaggio and Powell, 1983). Organizations adopt new practices not only for efficiency but to seek social and cultural legitimacy (Pasamar et al., 2025). The transition to regeneration can be analysed through institutional entrepreneurship lens, where powerful actors like the Armani Group challenge existing industrial norms to establish new "rules of the game". According to (Roszkowska-Menkes, 2023), this process is driven by three main isomorphic pressures:

- Coercive Pressures: Emerging environmental regulations and international climate agreements (e.g., the EU Green Deal) force companies to adopt regenerative standards to avoid legal and reputational risks (Pasamar et al., 2025).
- Normative Pressures: These arise from professionalization and the setting of new industry standards. As regenerative agriculture becomes defined by scientific and professional bodies, it creates a "logic of appropriateness" that influences corporate behavior (Gualandris et al., 2024).
- Mimetic Pressures: In environments of high uncertainty, such as the current climate crisis, organizations tend to model themselves after successful peers. By pioneering regenerative cotton in Apulia, Armani acts as a "template" for other luxury brands to follow (DiMaggio and Powell, 1983).

By integrating Institutional Theory with socio-ecological resilience, this paper argues that regeneration is more than a technical fix; it is a systemic shift that requires the alignment of institutional logics with ecological boundaries to foster long-term survival of both the industry and the territory (Gualandris and Madonna, 2025).

3 Methodology

This study adopts an exploratory qualitative case study methodology, drawing on established approaches in case study research (Eisenhardt, 1989; Stake, 2000; Yin, 2018). It focuses on Armani's Apulia Regenerative Cotton Project as a single, in-depth case. The research is designed as a single embedded case study, allowing for the analysis of multiple units of inquiry within the case. The single-case design is justified by the project's status as a critical, information-rich instance in which regenerative agricultural practices are implemented within a highly stressed socio-ecological context under the leadership of a globally influential corporate actor. Data collection relies on a triangulated approach combining (i) desk-based analysis and (ii) semi-structured interviews. Desk research involves systematically reviewing publicly available sources, including corporate reports, sustainability disclosures, institutional communications, and media coverage. This is complemented by primary data gathered through interviews with diverse stakeholders directly involved in the project, including farmers, local institutions, company representatives, and non-governmental organisations, enabling the exploration of multiple embedded perspectives within the case. The integration of secondary and primary data ensures methodological triangulation, enhancing the robustness and credibility of findings. This approach supports a comprehensive understanding of the initiative's governance, implementation dynamics, and local impacts, while enabling the development of a framework that is both theoretically grounded and empirically informed, with potential practical applicability beyond the case under study.

4 Results and Discussions

The Apulian Regenerative Cotton Project originated as a collaborative initiative involving multiple actors, including the Armani Group, the European Forest Institute (EFI), and the Italian Council for Agricultural Research and Economics (CREA), with the aim of exploring the feasibility of regenerative cotton production in Italy. The project was launched in 2023 as a pilot initiative designed to test agroforestry practices applied to cotton cultivation in a context characterized by environmental stress and declining soil quality. In the inception of the initiative, normative pressures played a central role, reflecting the growing expectations associated with sustainable production practices within the fashion industry. These pressures are also reflected in the involvement of scientific and institutional partners, which supported Armani in defining and legitimizing the business model through the adoption of regenerative agriculture practices. On the other hand, given the absence of established regenerative-based models in the fashion industry, mimetic pressures appear to be less relevant in the initial phase. Rather, the initiative led by Armani, as a leading player in the luxury fashion industry,

can be interpreted as an attempt to explore and define new business models in this field. Overall, in its early stage, the initiative can be seen as an attempt to translate emerging expectations into concrete actions, laying the foundations for the subsequent development and potential institutionalization of regenerative practices within the fashion industry. Following its initial configuration, the project moved into an implementation phase focused on the experimental application of regenerative agricultural practices. This phase was centred on the development of an agroforestry-based cotton production system, combining cotton cultivation with different tree species and applying regenerative techniques aimed at improving soil health, biodiversity, and water efficiency. The initiative was described by the partners as a "living laboratory", where scientific experimentation and practical implementation could progress simultaneously. This phase is characterized by the translation of pressures into concrete actions. Normative pressures continued to play a key role, as scientific partners contributed to defining appropriate methods for regenerative approaches. At the same time, coercive pressures may have played an indirect role, as recent frameworks and initiatives, such as TNFD and GRI, provide guidance and best practices for the implementation and disclosure of restoration-oriented initiatives. In this phase, the involvement of local agricultural firms also contributes to the emergence of new pressures, while at the same time laying the foundations for the institutionalization of the business model through the engagement of local actors. This highlights how practices are not simply adopted but actively constructed through interaction among different actors. In this regard, the introduction of monitoring activities serves to reinforce the legitimacy of the initiative by providing measurable outcomes for all partners. Overall, the implementation phase represents a critical step in the institutionalization process, as it allows emerging practices to be tested, refined, and gradually shaped into potentially replicable production practices. As the project progressed, the focus moved beyond the experimental phase toward the structuring of the supply chain for the use of regenerative cotton in garment production. This required Armani to involve suppliers and partners in defining production processes and packaging solutions. A central element of this phase is the introduction of certification and traceability mechanisms, which play a crucial role in supporting the credibility of the initiative. In particular, the project relies on the Regenagri certification scheme, which verifies the adoption of regenerative practices across the entire supply chain and ensures traceability from raw material to final product. The certification process involves the auditing of suppliers and production stages, requiring compliance with specific criteria related to soil health, biodiversity, and resource use. From an institutional theory perspective, this phase is primarily shaped by normative pressures. In the absence of fully established regulatory frameworks for regenerative agriculture, voluntary standards and certification schemes provide reference points that define what constitutes appropriate and legitimate practices. At the same time, certification and traceability mechanisms contribute to addressing the expectations of final consumers regarding product quality and origin. In this sense, the introduction of tools such as the digital product passport also responds to, and anticipates, emerging regulatory developments, suggesting the presence of potential coercive pressures. Overall, this phase represents a critical step in the institutionalization process, contributing to the establishment of mechanisms for communicating the initiative to external audiences and beginning to define shared standards and recognizable practices, which are essential for achieving legitimacy and enabling wider diffusion. Ultimately, the project led to the commercialization of the first batch of T-shirts made from regenerative cotton, distributed through selected Armani channels. This step demonstrated the technical and business viability of the initiative for Armani and its supply chain partners, potentially contributing in a decisive way to the legitimization and diffusion of business models based on regenerative approaches within the fashion industry. In this sense, the visibility of the project, supported by communication activities and the positioning of Armani as a leading luxury brand, increases the potential to trigger mimetic pressures among competitors. Due to the novelty of such initiatives, the Apulian Regenerative Cotton Project may become a reference point for other luxury fashion brands interested in developing business models based on regenerative approaches. The combination of scientific validation, engagement of local firms, and certification contributes to shaping a model that may be observed and potentially replicated by competitors or even lay the foundations for the institutionalization of regenerative practices. At the same time, this phase also highlights some limitations and challenges associated with the initiative. In particular, the positioning of the products within the luxury segment and the relatively small scale of production raise questions about the scalability of regenerative models. In addition, the leading role of Armani within the initiative and its position as a first mover raise questions regarding its role in defining normative pressures for partners and setting standards that may later be reinforced through mimetic isomorphism. Moreover, the lack of widely recognized regulatory frameworks highlights the current absence of strong formal coercive pressures and the prevalence of implicit ones, a condition that may evolve as these models become more widespread.

Overall, the longitudinal analysis of the Apulian Regenerative Cotton Project highlights a gradual process through which regenerative practices are progressively institutionalized both within the actors involved in the supply chain and toward the broader external environment. From an institutional perspective, the project shows

how multiple and interacting pressures contribute to shaping this process. Normative pressures, associated with scientific knowledge, voluntary standards, and sustainability expectations, play a central role in orienting actors' behaviours and practices. Mimetic pressures may emerge as the initiative gains visibility and begins to act as a reference point within the field, while coercive pressures remain largely limited to voluntary certification schemes, but may become more relevant as regulatory frameworks evolve. At the same time, the findings highlight that institutionalization is not a linear or automatic process. Rather, it depends on how actors interpret, negotiate, and translate institutional pressures into concrete practices within a specific context.

5 Implications

The journey of the Apulian Regenerative Cotton Project serves as more than just a corporate case study; it acts as a narrative of institutional transformation. By weaving together the threads of high fashion and soil science, this initiative reveals deep-seated implications for how we understand the future of the industry. From a theoretical standpoint, this research breathes new life into Institutional Theory by moving beyond the traditional view of companies as passive recipients of regulatory commands. Instead, we see the emergence of the brand as an institutional entrepreneur. In the absence of rigid laws, Armani does not wait for a mandate but collaborates with scientific pioneers to co-create a new logic of appropriateness. This suggests that in the early stages of systemic change, normative pressures, he shared belief in what is right or scientifically sound, are far more potent than formal rules. The project demonstrates that institutionalization is a social construction: by building a living laboratory, the actors involved are not just growing cotton; they are cultivating the very standards that will likely govern the fashion industry in the decade to come. The study thus proves that mimetic isomorphism is the final act of a successful pilot, where once the template is proven, the outlier becomes the benchmark for the entire field. For practitioners, the project offers a pragmatic blueprint for navigating the green transition while shattering the illusion that sustainability can be achieved through top-down directives alone. The primary takeaway for managers is that regenerative models live or die by their coordination. Success was not found in a simple purchase order, but in a deep, horizontal alignment between farmers, scientists, and certification bodies that bridged the gap between raw material and luxury product. Furthermore, by integrating tools like the Digital Product Passport and Regenagri certification today, firms can build a strategic regulatory buffer that allows them to stay ahead of looming EU mandates like the Carbon Removals Certification Framework. Perhaps the most vital practical lesson lies in the tension between luxury and volume. While the Armani model works for high-end garments, the true test for the industry will be translating these artisanal, territory-specific successes into global supply chain realities. Ultimately, the implications suggest that the future of fashion lies in radical transparency, where brands must become as invested in the health of the soil as they are in the aesthetics of the final garment.

6 Conclusions

The longitudinal analysis of the Apulian Regenerative Cotton Project provides a comprehensive understanding of how regenerative practices are nurtured, validated, and eventually institutionalized within the complex ecosystem of the luxury fashion industry. By applying Institutional Theory, this study clarifies that the transition to a regenerative business model is not merely a technical shift in agricultural inputs, but a profound social and organizational process. In the absence of mature regulatory frameworks, the interplay between normative expectations from scientific partners and the proactive entrepreneurship of a lead firm like Armani creates a new "logic of appropriateness". This logic serves as a precursor to broader industry standards, suggesting that institutionalization in fashion is increasingly a bottom-up phenomenon, fueled by cross-sector collaborations that bridge the gap between ecological restoration and commercial viability. From a practical standpoint, the study underscores that successful regenerative initiatives depend on the ability of a leading brand to act as a central coordinator. For firms aiming to remain competitive, the integration of certification schemes like Regenagri and traceability tools such as the Digital Product Passport is no longer optional but a strategic necessity to anticipate the risks of forthcoming European regulations. However, this research is not without its limitations. As an exploratory study centered on a single case in Southern Italy, the findings are inherently contextual. The "Armani model" may not be directly generalizable to smaller brands with less institutional weight or to different geographical regions with varying agricultural traditions. Furthermore, while we have mapped the early stages of institutionalization, it remains to be seen whether these practices will become deeply embedded industry norms or remain peripheral to the mass market. Future research should look toward comparative analyses across different market segments, such as fast fashion or outdoor apparel, to see how regenerative models survive under different price sensitivities. Additionally, as the regulatory environment evolves with the full implementation of the EU Carbon Removals Certification Framework, longitudinal studies

will be essential to assess the shift from voluntary to coercive pressures. There is also a significant opportunity to integrate quantitative data, such as Life Cycle Assessments, to evaluate if these “logics of appropriateness” translate into measurable ecological recovery and improved livelihoods for local farming communities.

7 Ethics declaration

The author(s) declare that the research presented in this paper did not involve human participants, animal subjects, or the collection of sensitive personal data. Consequently, ethical clearance from an institutional review board or ethics committee was not required for this study.

8 AI declaration

During the preparation of this work, the author(s) did not use any generative AI or AI-assisted tools in the writing process, data analysis, or development of the manuscript. The entire content of this paper is the original work of the author(s) and was produced without the assistance of artificial intelligence.

References

Beacham JD, Jackson P, Jaworski CC, et al. (2023) Contextualising farmer perspectives on regenerative agriculture: A post-productivist future? *Journal of Rural Studies* 102: 103100.

Brydges T (2021) Closing the loop on take, make, waste: Investigating circular economy practices in the Swedish fashion industry. *Journal of Cleaner Production* 293: 126245.

Colasante A, D’Adamo I, Desideri S, et al. (2025) Environmental Concerns in the Fashion Industry: A Twin Transition With the Digital Product Passport. *Business Strategy and the Environment* 34(7): 9242–9256.

DiMaggio PJ and Powell WW (1983) The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields. *American Sociological Review* 48(2): 147.

Eisenhardt KM (1989) Building theories from case study research. *Academy of management review* 14(4). 4. Academy of management Briarcliff Manor, NY 10510: 532–550.

Erzurumlu SS, Deets S, Nersessian D, et al. (2023) Strategic engagement of business with Sustainable Development Goals: A systems thinking approach. *Business Strategy and the Environment* 32(7): 4954–4969.

Gualandris J and Madonna A (2025) On the operational conditions enabling business-to-business waste synergies. *SSRN Electronic Journal*. Epub ahead of print 2025. DOI: 10.2139/ssrn.5220674.

Gualandris J, Branzei O, Wilhelm M, et al. (2024) Unchaining supply chains: Transformative leaps toward regenerating social–ecological systems. *Journal of Supply Chain Management* 60(1): 53–67.

Gugganig M (2025) Fixing sustainability through technoscience and diversity: The case of EU agriculture policy. *Environmental Science & Policy* 171: 104121.

Oliveira Duarte L, Vasques RA, Fonseca Filho H, et al. (2022) From fashion to farm: Green marketing innovation strategies in the Brazilian organic cotton ecosystem. *Journal of Cleaner Production* 360: 132196.

Pasamar S, Bornay-Barrachina M and Morales-Sánchez R (2025) Institutional pressures for sustainability: a triple bottom line approach. *European Journal of Management and Business Economics* 34(4): 460–484.

Peter John E and Mishra U (2023) A sustainable three-layer circular economic model with controllable waste, emission, and wastewater from the textile and fashion industry. *Journal of Cleaner Production* 388: 135642.

Roszkowska-Menkes M (2023) Institutional Theory. In: Idowu SO, Schmidpeter R, Capaldi N, et al. (eds) *Encyclopedia of Sustainable Management*. Cham: Springer International Publishing, pp. 1983–1990. Available at: https://link.springer.com/10.1007/978-3-031-25984-5_389 (accessed 30 April 2026).

Shahmohamadloo RS, Febria CM, Fraser EDG, et al. (2021) The sustainable agriculture imperative: A perspective on the need for an agrosystem approach to meet the United Nations Sustainable Development Goals by 2030. *Integrated Environmental Assessment and Management* 18(5): 1199–1205.

Shennan-Farpón Y, Vion-Loisel A, Van Soesbergen A, et al. (2025) Understanding regenerative agriculture in Europe: An analysis of academic literature, stakeholder perceptions and policy comparing Spain and the UK. *Environmental Science & Policy* 172: 104172.

Stake R (2000) Case studies (2000) Handbook of Qualitative Research. *NK Denzin and YS Lincoln (Eds.)*: 435–454.

Yin RK (2018) Case study research and applications: Design and methods. Thousand Oaks, CA: Sage. Epub ahead of print 2018.

Zucker LG (1987) Institutional theories of organization. *Annual review of sociology* 13. JSTOR: 443–464.

Integrating ESG Principles into Production Processes for Shaping Sustainable Industry

Mattia Arcuti, Vito Del Vecchio, Mariangela Lazoi, Giorgia Specchia
University of Salento, Lecce, Italy

Abstract

The integration of Environmental, Social, and Governance (ESG) principles into organizational processes has become a critical challenge for firms seeking to align sustainability commitments with operational decision-making, regulatory compliance, and long-term value creation. Although Enterprise Resource Planning (ERP) systems represent core infrastructures for data integration, process standardization, and organizational memory, their role in supporting ESG integration remains fragmented and underexplored. This paper investigates how ESG principles and data can be integrated into ERP systems from a knowledge management perspective, with particular attention to the management of green knowledge. The study adopts a Systematic Literature Review (SLR), performed on Scopus, and collect 507 results. The findings, based on the final resulting 29 papers, show that ESG–ERP integration address specific technologies, indicators, sectors, or implementation issues, while limited attention is paid to how ESG-related data, practices, and responsibilities are transformed into structured, reusable, and actionable green knowledge within ERP systems. To address this gap, the paper proposes a research agenda organized around three dimensions: human, technological, and organizational. The paper also develops a set of best practices for knowledge-based ESG–ERP integration. These practices suggest that organizations should start from existing ERP processes, define shared ESG data meanings, embed ESG controls into workflows and decision-making, and support users through training, communication, and continuous learning. The study contributes to the literature by reframing ESG–ERP integration as a knowledge management process and provides practical guidance for organizations seeking to integrate sustainability principles into ERP systems in a gradual, coherent, and non-disruptive way.

Keywords: ESG; ERP systems; Knowledge Management; Green Knowledge, ESG-ERP Integration.

Paper type: Academic Research Paper

1 Introduction

On July 24, 2025, humanity reached Earth Overshoot Day, the date on which global demand for ecological resources and services exceeded what the Earth can regenerate within the same year. From that day onward, humanity operated in a state of ecological deficit, drawing on natural capital and accumulating environmental pressure. Since the early 1970s, Earth Overshoot Day has generally occurred earlier in the calendar year, reflecting the growing imbalance between human demand and the planet's regenerative capacity. (*Earth Overshoot Day Home - #MoveTheDate*, n.d.).

Mitigating this crisis requires the involvement of multiple actors, including firms. In recent years, the United Nations 2030 Agenda and the articulation of the 17 Sustainable Development Goals (SDGs) have encouraged organizations to expand the scope of their organizational knowledge by systematically incorporating sustainability principles. Environmental, Social, and Governance (ESG) criteria have emerged as a key framework for disciplining and assessing corporate sustainability, risk exposure, and long-term value creation (Wang et al., 2024). As a result, companies are increasingly required to acquire, develop, and manage what can be defined as green knowledge (Sahoo et al., 2023). This concept does not refer to a narrow or standalone category, but rather to environment-related knowledge (data, skills, know-how, and ideas) that is acquired, stored, shared, applied, and created inside organizations to improve environmental performance, support green innovation, and advance sustainability goals (Ulhaq et al., 2024).

This perspective is consistent with the Knowledge-Based Theory proposed by (Grant, 1996), which conceptualizes the organization as a system for knowledge integration, in which technologies, as embodiments of knowledge, human interactions, and organizational routines jointly contribute to value creation. If knowledge is inherently distributed and dynamic, its effective management requires technological infrastructures that spread beyond information storage and can support integration, coordination, and organizational memory (Duffy, 2000). In this context, Enterprise Resource Planning (ERP) systems emerge as potential enablers of organizational knowledge (Saide & Mahendrawathi, 2015). ERP systems can be interpreted as repositories of organizational knowledge (Palanisamy, 2008), as they enable firms to store and institutionalize so-called canonical organizational processes (Vandaie, 2008). Through process codification, ERP systems embed a

significant portion of organizational memory, making knowledge accessible, replicable, and governable over time (Vandaie, 2008).

The integration of ESG principles into ERP systems, however, raises specific knowledge management challenges. Drawing on Nonaka's theoretical framework, a distinction can be made between tacit and explicit knowledge (Nonaka & Takeuchi, 1995), ESG integration can be understood as a knowledge conversion process through which dispersed, heterogeneous, and often tacit sustainability-related knowledge must be transformed into explicit, structured, and reusable organizational knowledge. At the operational level, this process is particularly complex because firms deal with a heterogeneous mix of data, information, practices, regulatory requirements, and expertise that is difficult to disentangle and standardize (Vandaie, 2008).

This misalignment between the strategic relevance of ESG and its operational integration also emerge from preliminary focus groups conducted with C-level executives from a global IT consulting company. The discussions highlighted that, although digital technologies and sustainability are widely acknowledged as strategic levers, the integration of ESG data and principles into ERP systems remains fragmented, underexplored, and often deprioritized in implementation projects.

Against this background, the present study aims to investigate how ESG principles can be integrated into ERP systems from a knowledge management perspective. More specifically, the paper addresses the following research question: *How can ESG principles and data be integrated into ERP systems to support the management of green knowledge?*

To answer this question, the study adopts a qualitative and exploratory research design based on a Systematic Literature Review (SLR) conducted according to the consolidated protocol proposed by (Corallo et al., 2023). The findings of the study are articulated around two main contributions. First, the paper proposes a research agenda aimed at guiding future investigations on the integration of sustainability, ERP systems, and knowledge management. Second, it identifies a set of guiding principles intended to support organizations in the integration of ESG data into ERP systems.

The remainder of the paper is organized as follows. Section 2 describes the research methodology and the SLR protocol. Section 3 presents the findings of the literature review. Section 4 outlines the research agenda and proposes a set of best practices. Finally, Section 5 concludes the paper by summarizing its theoretical and managerial implications, limitations, and future research opportunities.

2 Methodology

This study adopts a qualitative and exploratory research design based on a Systematic Literature Review (SLR). The review was conducted following the protocol proposed by Corallo et al. (2023), which structures the process into four main phases: review planning, search execution, documents analysis, and results reporting. This approach was selected because it provides a transparent and replicable procedure for identifying, selecting, and synthesizing scientific contributions on an emerging and interdisciplinary topic (Tranfield et al., 2003).

In the review planning phase, the research objective was defined in line with the research question of the study. The search strategy was developed to capture interdisciplinary contributions. The query combined four groups of keywords. The first group referred to ERP systems, including both the full expression "Enterprise Resource Planning" and the acronym "ERP". The second group captured ESG and sustainability-related concepts, including "ESG", "Environmental Social Governance", "sustainab*", "sustainable reporting", "sustainability reporting", and "green ERP". The third group focused on data and information systems, including "data", "data management", "data architecture", "data model*", and "information system*". Finally, the fourth group included terms related to implementation and managerial processes, namely "integrat*", "implement*", "report*", and "monitor*". The final search string was applied to titles, abstracts, and keywords as follows:

TITLE-ABS-KEY(("Enterprise Resource Planning" OR ERP) AND (ESG OR "Environmental Social Governance" OR sustainab OR "sustainable reporting" OR "sustainability reporting" OR "green ERP") AND ("data" OR "data management" OR "data architecture" OR "data model*" OR "information system*") AND (integrat* OR implement* OR report* OR monitor*)).*

In the search execution phase, the bibliographic search was conducted in Scopus in October 2025. Scopus was selected because of its broad coverage of peer-reviewed academic literature. The initial search returned 507 documents. The results were then refined by limiting the sample to English-language publications, open-access documents, and journal articles or conference papers. The open-access filter was applied to ensure the full availability of papers for complete reading and analysis, while the document-type filter was used to include both consolidated academic contributions and emerging research trends. After these refinements, the resulting sample consisted of 105 papers.

Table 1: Screening process.

Initial Scopus results	English papers	Document type refinement	Open-access refinement	Title-abstract reading	Full-text reading
507	501	402	105	75	29

In the documents analysis phase, the titles and abstracts of the 105 papers were screened to assess their relevance to the research objective. This screening led to the exclusion of 30 papers that were not aligned with the focus of the study, leaving 75 papers for full-text assessment. The full-text analysis was conducted using one inclusion criterion and three exclusion criteria, defined ex ante. The inclusion criterion was:

- IC1: the paper addresses the relationship between ERP systems and the integration of ESG reporting data or sustainability-related data.

The exclusion criteria were:

- EC2: the paper does not address either ERP systems or ESG/sustainability reporting criteria;
- EC3: the paper focuses only on ERP or enterprise technologies, without considering ESG or sustainability-related data integration;
- EC4: the paper focuses only on ESG or sustainability issues, without considering their connection with ERP systems or enterprise information systems.

In Table 1 is resume the entire screening process, while in the Table 2 is reported the distribution of the paper according to the inclusion and exclusion criteria.

Table 2: Distribution of papers within the inclusion and exclusion criteria.

Papers IC1	Papers EC2	Papers EC3	Papers EC4
29	24	21	1

In the results reporting phase, the selected contributions were synthesized through a thematic and comparative analysis. This process made it possible to identify recurring patterns, operational barriers, methodological differences, and research gaps in the literature. Particular attention was paid to how ERP systems are conceptualized as infrastructures for data integration, process codification, sustainability reporting, regulatory compliance, and knowledge-based decision-making. The emerging themes provided the basis for developing the guiding principles for ESG integration into ERP systems and the research agenda presented in the following sections.

3 State of Art

Overall, the reviewed literature converges in recognizing that ERP systems are increasingly evolving beyond their traditional role of supporting resource planning and transactional processes. They are progressively conceptualized as integrated platforms for monitoring, measuring, and reporting sustainability performance. This evolution is consistent with the emerging notion of Sustainable ERP (S-ERP), according to which enterprise information systems should embed ESG-related metrics, data flows, and processes within core organizational routines rather than treating sustainability as an external reporting activity (Adriansyah et al., 2022; Alzahmi et al., 2025)

A first stream of contributions emphasizes the role of ERP systems as infrastructures for ESG data integration across business functions and supply chains. From this perspective, ERP systems support the collection, standardization, and sharing of sustainability-related data, enabling organizations to connect environmental and social indicators with operational processes. Empirical studies on ERP and Green Supply Chain Management show that this integration can improve both operational and environmental performance by supporting green procurement, environmental monitoring, production optimization, and the reduction of waste, energy consumption, and emissions (Rizki et al., 2022). Similarly, other studies highlight that ERP-enabled sustainability practices may contribute simultaneously to economic growth and environmental improvement, suggesting that ERP systems can act as mediating infrastructures between business efficiency and sustainability objectives (Anaya et al., 2023; Anaya & Qutaishat, 2022).

A second trajectory concerns the integration of ESG data along the product and supply chain life cycle. In this area, the literature highlights the need to connect ERP systems with Product Lifecycle Management (PLM), Life Cycle Assessment (LCA) tools, Warehouse Management Systems (WMS), and Manufacturing Execution Systems (MES). Such integration is considered essential to overcome data fragmentation, improve the completeness and reliability of sustainability information, and support more accurate ESG reporting (Felder et al., 2025; Loy et al., 2024; Rosich et al., 2013; Tong et al., 2023). This stream of research shows that ESG–ERP integration cannot be

reduced to the inclusion of additional reporting fields within ERP modules. Rather, it requires the alignment of multiple operational systems and the construction of consistent data flows across the entire value chain.

A third group of studies focuses on the measurement and evaluation of sustainability performance through indicators, KPIs, and decision-support models. These contributions address the methodological complexity of selecting, weighting, and operationalizing ESG indicators within enterprise systems. Approaches based on multi-criteria decision-making methods, such as Fuzzy Delphi, Analytic Hierarchy Process (AHP), ANP, TOPSIS, and MCGP, are used to support the selection of ESG indicators, suppliers, strategies, and knowledge-transfer mechanisms (Cano et al., 2025a, 2025b; Cosma et al., 2024; Lacurezeanu et al., 2021; Patalas-Maliszewska et al., 2022; Susanto et al., 2023). These studies suggest that ESG integration into ERP systems requires not only data availability, but also structured evaluation logics capable of translating sustainability goals into measurable and comparable decision criteria.

Another relevant line of research investigates the relationship between sustainability and economic efficiency, particularly in complex and resource-intensive sectors such as construction and manufacturing. In these contexts, ERP systems are presented as tools for integrating cost control, resource allocation, planning, and environmental performance. The literature shows that ERP-enabled sustainability management can support the balancing of economic and environmental objectives by improving visibility over resource use, process inefficiencies, and operational trade-offs (Mandičák et al., 2022; Tsai, 2023). This perspective reinforces the idea that ESG–ERP integration is not only a compliance-oriented activity, but also a mechanism for improving organizational efficiency and strategic decision-making.

From a technological standpoint, the literature also points to the increasing automation and intelligence of ESG-related ERP processes. Process mining has been proposed as a means to support automated ESG compliance control, while digital reporting tools and operational-data reinterpretation approaches contribute to transforming sustainability reporting from a manual and retrospective activity into a more continuous and data-driven process (Anjaria, 2024; Dudek & Kulej-Dudek, 2024; Schreiber, 2020). These contributions indicate a shift from static ESG reporting toward dynamic monitoring, where ERP systems interact with analytics, automation, and intelligent technologies to improve transparency, traceability, and responsiveness.

Alongside technological aspects, the literature stresses the relevance of organizational and human factors. ESG-ERP integration depends on the alignment between technical systems, organizational structures, governance mechanisms, data quality, and user acceptance. Studies on ERP implementation and sustainability-oriented information systems show that strategic alignment, top management support, process ownership, and data governance are critical conditions for effective integration (Hadi et al., 2018; Huang et al., 2019). Moreover, technology acceptance models highlight that perceived usefulness, user expectations, and behavioral intention influence the extent to which ERP systems are effectively adopted and used to generate sustainability-related value (Junaidi et al., 2024; Marsintauli & Sari, 2023).

A final and particularly relevant stream concerns the knowledge management dimension of ESG-ERP integration. From this perspective, ERP systems are not only repositories of sustainability data, but infrastructures that support the codification, transfer, and reuse of green knowledge. Knowledge Management strengthens the role of ERP systems in supporting sustainable decision-making, especially when sustainability-related information is dispersed across departments, processes, and external stakeholders (Yahya & Khalefa, 2021). More advanced approaches, such as the integration of ERP systems with intelligent systems, digital twins, and hybrid knowledge-based systems, further show how data, models, and expert knowledge can be combined to support more effective and sustainable decisions within Industry 5.0 contexts (Traini et al., 2024). This stream is particularly important for the present study because it reframes ESG-ERP integration as a process of transforming fragmented sustainability knowledge into structured, accessible, and actionable organizational knowledge.

Taken together, the reviewed literature shows that ESG–ERP integration is evolving from a reporting-oriented and technical issue into a broader socio-technical challenge involving data integration, process codification, sustainability measurement, governance, and knowledge management. However, the field remains fragmented: existing studies often focus on specific technologies, indicators, sectors, or implementation issues, while limited attention is paid to how dispersed ESG-related data and practices are transformed into structured, reusable, and actionable green knowledge within ERP systems.

4 Best Practices and Research Agenda

The gap, highlighted in reviewing analysis, suggests that ESG-ERP integration should be understood as a knowledge management process through which organizations codify sustainability knowledge, assign responsibilities, ensure data quality and traceability, and embed ESG principles into operational and strategic

decision-making. This supports the need to identify guiding principles for ESG–ERP integration and to define a research agenda capable of connecting sustainability, enterprise systems, and green knowledge management. The fragmentation of the studies suggests that future research should move beyond the technical question of how ESG indicators can be added to ERP platforms and investigate the human, technological, and organizational conditions that enable ESG-ERP integration as a knowledge management process. Accordingly, the following research questions are proposed.

Table 3: Proposed research agenda.

Dimensions	Research questions
Human	RQ1. Which green skills, competencies, and behaviors enable the effective integration of ESG principles into ERP systems?
	RQ2. How does green knowledge influence the interpretation and use of information generated by ERP systems?
	RQ3. How is the green transition of ERP systems perceived by different organizational actors, such as top management, middle managers, and operational users?
Technological	RQ4. Does the green transition of ERP systems require a radical redesign of existing platforms, or can it be achieved through incremental adaptation?
	RQ5. How can ERP systems remain aligned with the continuous evolution of ESG requirements, regulations, and indicators?
	RQ6. How can ERP systems integrate external, heterogeneous, and often unstructured ESG data while ensuring reliability, traceability, and timely updating?
Organizational	RQ7. Can Sustainable ERP systems be introduced as substitutes for traditional ERP systems, or do they require structured change management processes?
	RQ8. How can ESG-ERP integration support the transformation of green knowledge into ordinary organizational knowledge and routines?
	RQ9. What are the short-, medium-, and long-term effects of ESG–ERP integration on organizational performance, costs, resource allocation, compliance, and decision-making?

From an organizational point of view, firms need actionable best practices to translate ESG objectives into ERP-enabled processes. In the knowledge management perspective, the objective is not only to add ESG indicators to existing ERP modules, but to transform dispersed sustainability-related data, practices, and expertise into codified, shared, and actionable green knowledge. To reduce organizational disruption and support adoption, the following best practices are proposed.

- **BP1. Start from existing ERP systems and define shared ESG data meanings:** ESG integration should begin by mapping current ERP workflows and identifying where ESG data are already generated or used. Organizations should then translate tacit and dispersed ESG knowledge into shared data definitions, KPI meanings, and responsibilities. Linking ESG requirements to existing processes, supported by data dictionaries and common glossaries, reduces duplication, avoids parallel reporting systems, limits ambiguity, and supports user acceptance across functions.
- **BP2. Embed ESG controls into workflows and decision-making:** ESG controls should be integrated into daily ERP routines through predefined fields, alerts, approval rules, automated data capture, and audit trails. At the same time, ERP dashboards, reports, and budgeting tools should connect ESG indicators to operational and strategic decisions. This makes ESG knowledge traceable and actionable without creating an additional bureaucratic layer.
- **BP3. Support users and institutionalize continuous learning:** Users should understand both how to manage ESG data in ERP systems and why these data matter. Training, communication, workshops, and feedback loops can improve adoption, usability, and the capture of new tacit knowledge. At the same time, ERP systems should support the periodic revision of ESG indicators, procedures, responsibilities, and reporting structures, allowing integration to evolve with regulations, standards, technologies, and stakeholder expectations.

5 Discussions and Conclusions

This study contributes to the literature by reframing ESG–ERP integration as a knowledge management challenge rather than as a merely technical or reporting-oriented issue. The systematic literature review shows that existing research has mainly addressed ESG–ERP integration through specific lenses, such as sustainability reporting, data integration, KPI measurement, supply chain monitoring, compliance automation, or digital transformation. While these perspectives provide valuable insights, they often remain fragmented and do not

fully explain how ESG-related data, practices, responsibilities, and expertise can be transformed into structured, reusable, and actionable green knowledge within ERP systems. By adopting a knowledge management perspective, this paper highlights that ERP systems can act not only as repositories of organizational data, but also as infrastructures for codifying, sharing, and institutionalizing sustainability-related knowledge across organizational functions.

From a theoretical standpoint, the paper advances the debate on ERP systems, ESG integration, and knowledge management in three ways. First, it connects the literature on enterprise systems with the concept of green knowledge, showing that sustainability integration requires the transformation of dispersed and often tacit ESG knowledge into explicit and systematized organizational knowledge. Second, it extends the interpretation of ERP systems beyond their transactional and process-standardization role, emphasizing their potential contribution to organizational learning, ESG traceability, accountability, and decision-making. Third, it proposes a research agenda organized around the human, technological, and organizational dimensions of ESG–ERP integration. This agenda helps structure future research by identifying open questions related to green skills and user perceptions, the technological adaptability of ERP platforms, the integration of external and unstructured ESG data, and the organizational conditions required to embed sustainability into ordinary routines.

The managerial implications are equally relevant. The findings suggest that organizations should not approach ESG–ERP integration as the simple addition of sustainability indicators to existing modules. Instead, they should manage it as a gradual socio-technical transformation. This approach can reduce human resistance, avoid technological disruption, and prevent the creation of parallel reporting systems. It also helps managers clarify data ownership, improve data quality, and connect ESG information to operational and strategic decisions. In practical terms, ERP systems can support environmental management by linking resource consumption and emissions to production, logistics, and procurement processes; social sustainability by integrating workforce, safety, and supplier-related indicators; and governance by embedding compliance controls, ESG-related cost centers, audit trails, and accountability mechanisms into daily routines. The study also provides implications for consultants, ERP vendors, and sustainability managers. Consultants can use the proposed best practices as a roadmap for designing ESG–ERP integration projects that are less disruptive and more aligned with organizational knowledge flows. ERP vendors can benefit from developing functionalities that support ESG data governance, interoperability with external sources, traceability, and dynamic updating of sustainability indicators. Sustainability managers, on the other hand, can use ERP systems to move ESG management from a periodic reporting activity to an embedded decision-support process. In this sense, the integration of ESG principles into ERP systems may help organizations shift from compliance-oriented sustainability to knowledge-based sustainability management.

This study is subject to some limitations due to its exploratory and conceptual nature. First, the proposed research agenda and best practices are derived from the literature and from preliminary insights used to frame the research problem, but they require further empirical validation. Future research could test them through case studies, surveys, Delphi studies, or action research involving firms implementing ESG functionalities within ERP systems. Second, the study treats ESG-ERP integration as a general organizational challenge, while the relevance of specific ESG indicators, data flows, and regulatory requirements may vary significantly across industries. Future research could therefore investigate sector-specific configurations, for example in manufacturing, logistics, construction, energy, or services. Finally, the proposed best practices provide a general roadmap, but they do not define a maturity model or implementation sequence. Future studies could develop staged models to assess how organizations progressively move from fragmented ESG reporting to fully embedded, knowledge-based ESG management within ERP systems.

6 Ethics declaration

Ethical clearance was not required for the SLR, as it is based on publicly available academic sources.

7 AI declaration

AI-assisted tools were used during the preparation of this paper to support language editing. The authors critically reviewed, revised, and validated all AI-assisted outputs and remain fully responsible for the content, accuracy, and integrity of the submitted paper.

References

Adriansyah, A. K., Ridwan, A. Y., & Hediyanto, U. Y. K. S. (2022). Designing and ERP System: A Sustainability Approach. *International Journal on Informatics Visualization*, *6*(2–2), 498–505. Scopus. https://doi.org/10.30630/joiv.6.2-2.1117

Alzahmi, W., Al-Assaf, K., Alshaikh, R., & Bahroun, Z. (2025). Towards Sustainable ERP Systems: Emerging Trends, Challenges, and Future Pathways. *Management Systems in Production Engineering*, *33*(1), 24–38. Scopus. https://doi.org/10.2478/mspe-2025-0003

Anaya, L., Flak, L., & Abushakra, A. (2023). Realizing Sustainable Value from ERP Systems Implementation. *Sustainability (Switzerland)*, *15*(7). Scopus. https://doi.org/10.3390/su15075783

Anaya, L., & Qutaishat, F. (2022). ERP systems drive businesses towards growth and sustainability. In Cruz-Cunha M.M. & Mateus-Coelho N. (Eds.), *Procedia Comput. Sci.* (Vol. 204, pp. 854–861). Elsevier B.V. Scopus. https://doi.org/10.1016/j.procs.2022.08.103

Anjaria, K. (2024). Enhancing sustainability integration in Sustainable Enterprise Resource Planning (S-ERP) system: Application of Transaction Cost Theory and case study analysis. *International Journal of Information Management Data Insights*, *4*(2). Scopus. https://doi.org/10.1016/j.jjimei.2024.100243

Cano, A. K. S., Ali, M. B., Rauch, E., & Matt, D. T. (2025a). Environmental Performance Measurement in the Precast Concrete Sector Using a Fuzzy Logic-Based Stakeholder-Driven Indicator Selection. *Results in Engineering*, 106667.

Cano, A. K. S., Ali, M. B., Rauch, E., & Matt, D. T. (2025b). Strategic Selection of Environmental Indicators in Precast Concrete Industry Using a Fuzzy Delphi Analytic Hierarchy Process. *IFAC-PapersOnLine*, *59*(10), 879–884.

Corallo, A., Crespino, A. M., Vecchio, V. D., Lazoi, M., & Marra, M. (2023). Understanding and Defining Dark Data for the Manufacturing Industry. *IEEE Transactions on Engineering Management*, *70*(2), 700–712. https://doi.org/10.1109/TEM.2021.3051981

Cosma, A., Conte, R., Solina, V., & Ambrogio, G. (2024). Design of KPIs for evaluating the environmental impact of warehouse operations: A case study. In Longo F., Shen W., & Padovano A. (Eds.), *Procedia Comput. Sci.* (Vol. 232, pp. 2701–2708). Elsevier B.V. Scopus. https://doi.org/10.1016/j.procs.2024.02.087

Dudek, D., & Kulej-Dudek, E. (2024). Modern technologies in ESG reporting—Evidence from Polish enterprises. In Flearmoy J. (Ed.), *Procedia Comput. Sci.* (Vol. 246, Issue C, pp. 5359–5367). Elsevier B.V. Scopus. https://doi.org/10.1016/j.procs.2024.09.659

Duffy, J. (2000). The KM technology infrastructure. *Information Management Journal*, *34*(2), 62–66.

Earth Overshoot Day home—#MoveTheDate. (n.d.). Retrieved April 30, 2026, from https://overshoot.footprintnetwork.org/

Felder, M., Bataleblu, A. A., Grünbacher, G., & Rauch, E. (2025). Development of an ERP-Integrated Direct Routing and Way-Point Routing for Increasing Automation of LCAs in Supply Chains. In Solina V., Longo F., & Romero D. (Eds.), *Procedia Comput. Sci.* (Vol. 253, pp. 2674–2683). Elsevier B.V. Scopus. https://doi.org/10.1016/j.procs.2025.01.327

Grant, R. M. (1996). Toward a knowledge-based theory of the firm. *Strategic Management Journal*, *17*(S2), 109–122. https://doi.org/10.1002/smj.4250171110

Hadi, A.-A., Alnoor, A., & Abdullah, H. O. (2018). Socio-technical approach, decision-making environment, and sustainable performance: Role of ERP systems. *Interdisciplinary Journal of Information, Knowledge, and Management*, *13*, 397–415.

Huang, S. Y., Chiu, A. A., Chao, P. C., & Arniati, A. (2019). Critical success factors in implementing enterprise resource planning systems for sustainable corporations. *Sustainability (Switzerland)*, *11*(23). Scopus. https://doi.org/10.3390/su11236785

Junaidi, K., Rahmanta, & Wibowo, R. (2024). *The influence of perceived usefulness in the technology acceptance model on the success of ERP implementation to support sustainability in North Sumatra's palm oil companies*. *1413*(1), 012096.

Lacurezeanu, R., Chis, A., & Bresfelean, V. P. (2021). Integrated management solution for a sustainable SME-selection proposal using AHP. *Sustainability (Switzerland)*, *13*(19). Scopus. https://doi.org/10.3390/su131910616

Loy, R., Britton, L. L., & Malone, T. (2024). Software solutions in agri-food supply chains: A current view for sustainability reporting. *International Food and Agribusiness Management Review*, *27*(4), 729–744. Scopus. https://doi.org/10.22434/IFAM.1085

Mandičák, T., Spišáková, M., Mésároš, P., & Kozlovská, M. (2022). Design of Economic Sustainability Supported by Enterprise Resource Planning Systems in Architecture, Engineering, and Construction. *Buildings*, *12*(12). Scopus. https://doi.org/10.3390/buildings12122241

Marsintauli, F., & Sari, M. (2023). Evaluation of ERP Oracle NetSuite Implementation Using the Unified Theory of Acceptance and Use of Technology (UTAUT) Model to Create a Sustainable Business. In Mursitama T.N., Noerlina null, Utama D.N., & Abrori S.A. (Eds.), *E3S Web Conf.* (Vol. 426). EDP Sciences. Scopus. https://doi.org/10.1051/e3sconf/202342602037

Nonaka, I., & Takeuchi, H. (1995). *The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation*. Oxford University Press. https://doi.org/10.1093/oso/9780195092691.001.0001

Palanisamy, R. (2008). Organizational Culture and Knowledge Management in ERP Implementation: An Empirical Study. *Journal of Computer Information Systems*, *48*(2), 100–120. https://doi.org/10.1080/08874417.2008.11646013

Patalas-Maliszewska, J., Kłos, S., & Dostatni, E. (2022). Sustainable Business Model integrated with the Enterprise Resource Planning system: SBM-ERP. *Bulletin of the Polish Academy of Sciences: Technical Sciences*, *70*(6). Scopus. https://doi.org/10.24425/bpasts.2022.143829

Rizki, A. F., Murwaningsari, E., & Sudibyo, Y. A. (2022). Does Green Supply Chain Management Improve Sustainable Performance? *International Journal of Energy Economics and Policy*, *12*(6), 323–331. Scopus. https://doi.org/10.32479/ijeep.13648

Rosich, M. B., Le Duigou, J., & Bosch-Mauchand, M. (2013). Implementing sustainable supply chain in PLM. *IFIP Advances in Information and Communication Technology*, *398*(PART 2), 168–175. Scopus. https://doi.org/10.1007/978-3-642-40361-3_22

Sahoo, S., Kumar, A., & Upadhyay, A. (2023). How do green knowledge management and green technology innovation impact corporate environmental performance? Understanding the role of green knowledge acquisition. *Business Strategy and the Environment*, *32*(1), 551–569. https://doi.org/10.1002/bse.3160

Saide, & Mahendrawathi, E. R. (2015). Knowledge Management Support for Enterprise Resource Planning Implementation. *Procedia Computer Science*, *The Third Information Systems International Conference 2015*, *72*, 613–621. https://doi.org/10.1016/j.procs.2015.12.170

Schreiber, C. (2020). Automated Sustainability Compliance Checking Using Process Mining and Formal Logic. *ACM Int. Conf. Proc. Ser.*, 181–184. Scopus. https://doi.org/10.1145/3401335.3401355

Susanto, E., Othman, N. A., Praseptiawan, M., Imran, A., & Hendriana, Y. (2023). A Network-Multi-Criteria Decision Analysis with Super Decision for Vendor Selection in the Automotive Component Industry. *Malaysian Journal of Consumer and Family Economics*, *31*, 125–154. Scopus. https://doi.org/10.60016/majcafe.v31.05

Tong, Q., Ming, X., & Zhang, X. (2023). Construction of Sustainable Digital Factory for Automated Warehouse Based on Integration of ERP and WMS. *Sustainability (Switzerland)*, *15*(2). Scopus. https://doi.org/10.3390/su15021022

Traini, E., Antal, G., Bruno, G., De Maddis, M., Lombardi, F., Panza, L., & Spena, P. R. (2024). Hybrid knowledge based system supporting Digital Twins in the Industry 5.0. In Longo F., Shen W., & Padovano A. (Eds.), *Procedia Comput. Sci.* (Vol. 232, pp. 1471–1480). Elsevier B.V. Scopus. https://doi.org/10.1016/j.procs.2024.01.145

Tranfield, D., Denyer, D., & Smart, P. (2003). Towards a Methodology for Developing Evidence-Informed Management Knowledge by Means of Systematic Review. *British Journal of Management*, *14*(3), 207–222. https://doi.org/10.1111/1467-8551.00375

Tsai, W.-H. (2023). Balancing Profit and Environmental Sustainability with Carbon Emissions Management and Industry 4.0 Technologies. *Energies*, *16*(17). Scopus. https://doi.org/10.3390/en16176175

Ulhaq, I., Nayak, R., George, M., Nguyen, H., & Quang, H. (2024). Green knowledge management: A bibliometric analysis, research trends and future directions. *VINE Journal of Information and Knowledge Management Systems*. https://doi.org/10.1108/VJIKMS-02-2024-0069

Vandaie, R. (2008). The role of organizational knowledge management in successful ERP implementation projects. *Knowledge-Based Systems*, *21*(8), 920–926. https://doi.org/10.1016/j.knosys.2008.04.001

Wang, N., Pan, H., Feng, Y., & Du, S. (2024). How do ESG practices create value for businesses? Research review and prospects. *Sustainability Accounting, Management and Policy Journal*, *15*(5), 1155–1177. https://doi.org/10.1108/SAMPJ-12-2021-0515

Yahya, S. M., & Khalefa, M. S. (2021). Knowledge management model to enhance Enterprise Resource Planning system in Green software development process. In Obaid A.J., Abdullatif F.A., Tolephih M.H., Hasan H.A., Albermany S., & Ahmed H.A. (Eds.), *J. Phys. Conf. Ser.* (Vol. 1879, Issue 2). IOP Publishing Ltd. Scopus. https://doi.org/10.1088/1742-6596/1879/2/022082

Explainable Artificial Intelligence as an Organisational Learning Lever: A Theoretical-Methodological Framework for Professional Upskilling in AI-Augmented Workplaces

Raimondo Fanale
Universitas Mercatorum, Rome, Italy

Abstract

The integration of Artificial Intelligence (AI) into professional practice is reshaping how knowledge is created, validated, and transmitted within organisations. While AI-driven decision support systems offer considerable analytical power, many rely on deep learning architectures that operate as black boxes, producing outputs that are difficult for human users to understand or critically evaluate. Explainable Artificial Intelligence (XAI) has emerged as a response to this opacity, yet its role has been framed predominantly in terms of technical transparency and user trust, underexploring its potential as a lever for professional development and organisational learning. This paper makes a theoretical-methodological contribution by introducing the XAI-OLL (Explainable Artificial Intelligence as an Organisational Learning Lever) framework, which reconceptualises XAI-generated explanations not merely as transparency mechanisms but as cognitive and organisational artefacts capable of activating processes of professional upskilling and knowledge integration. The framework proposes a sequential cognitive mechanism in which exposure to XAI explanations stimulates reflective reasoning — the deliberate comparison between algorithmic outputs and one's own judgment — which in turn fosters interpretive competence, defined as the capacity to critically engage with AI-generated evidence. This pathway is conditioned by explanation quality: imperfect or misleading explanations can attenuate or reverse the effect on calibrated trust. Interpretive competence enables calibrated trust, a proportionate form of reliance grounded in evidence quality rather than automation deference. Calibrated trust, in turn, facilitates the integration of AI-derived insights into shared professional knowledge, supporting organisational learning at the collective level. The XAI-OLL framework generates five testable propositions, operationalised through a pilot quasi-experimental protocol in AI-supported histopathology decision-making. The paper contributes to knowledge management, organisational learning, and HR transformation literature by positioning AI interpretive literacy — the competence to engage critically with XAI outputs — as a new dimension of professional capability in AI-augmented workplaces. Practical implications address HR professionals as active designers of XAI-enabled upskilling pathways, governance conditions, and organisational learning infrastructures.

Keywords: Explainable AI; organisational learning; professional upskilling; calibrated trust; XAI-OLL; HR transformation; knowledge management

Paper type: Theoretical and Methodological Contribution

1 Introduction

Artificial Intelligence (AI) is increasingly embedded in complex professional decision-making contexts, from medical diagnosis to financial risk assessment and industrial inspection. In these domains, deep learning systems generate predictions that carry significant epistemic weight, yet their internal logic remains largely opaque to the humans who rely on them. This opacity creates a fundamental tension for organisational learning: scholarship in this field has long linked individual and collective knowledge development to cycles of sense-making, deliberate reflection, and knowledge sharing (Nonaka and Takeuchi, 1995; Schön, 1983). When AI outputs arrive as unexplained predictions, they disrupt these cycles — users cannot reflect on criteria, compare interpretations, or integrate new insights into shared professional knowledge frameworks.

Explainable Artificial Intelligence (XAI) — the field of methods designed to make algorithmic decisions interpretable by human users (Arrieta et al., 2020) — has emerged as a direct response to this challenge. Yet despite rapid technical progress, the organisational and learning dimensions of XAI remain empirically underdeveloped. A systematic review of the XAI literature reveals that fewer than 0.7% of published papers validate explanations with human participants in real-world tasks (Chromik and Schubert, 2025). This gap is not merely a methodological shortcoming: it reflects a conceptual blind spot in which XAI is treated as a problem of output generation rather than a question of how professionals learn, adapt, and create knowledge in the presence of AI.

The central research question of this paper is: through what cognitive and organisational mechanisms do XAI explanations activate professional learning and upskilling, and what conditions determine whether those mechanisms succeed or fail? This is explicitly a question of professional development — not merely individual

decision accuracy — and it requires a framework that spans cognitive, organisational, and HR transformation perspectives.

This paper addresses that gap by proposing XAI-OLL (Explainable Artificial Intelligence as an Organisational Learning Lever), a theoretical-methodological framework that repositions AI-generated explanations as cognitive and organisational artefacts capable of triggering professional upskilling and collective knowledge integration. Drawing on dual-process theory (Kahneman, 2011), reflective practice (Schön, 1983), XAI metacognition research (von Zahn et al., 2024), trust calibration (Lee and See, 2004; Parasuraman et al., 1997), and the SECI model (Nonaka and Takeuchi, 1995), XAI-OLL proposes a sequential cognitive mechanism through which explainability activates reflective reasoning, builds interpretive competence, calibrates trust, and ultimately facilitates organisational knowledge integration. The framework is accompanied by a pilot quasi-experimental protocol designed to operationalise and validate its constructs in AI-supported histopathology decision-making.

2 Literature Review

2.1 Explainable Artificial Intelligence: Scope and Limitations

XAI encompasses methods designed to render the decisions of machine learning models interpretable by human users (Doshi-Velez and Kim, 2017). Arrieta et al. (2020) provide the foundational taxonomy, distinguishing intrinsically interpretable models from post-hoc explanation methods applied to opaque architectures. Post-hoc methods include local attribution techniques — LIME (Ribeiro et al., 2016) and SHAP (Lundberg and Lee, 2017) — that estimate feature contributions to individual predictions, and visual attribution methods — Grad-CAM (Selvaraju et al., 2017) and saliency maps — that highlight image regions most influential to a model's output. Longo et al. (2024) identify the human-centred and organisational dimensions of XAI as a priority research frontier, arguing for a shift from explanation as a technical property to explanation as a social and cognitive event.

However, empirical evidence on whether XAI explanations actually improve human decision-making and learning is mixed and frequently negative. Humer et al. (2024) demonstrate that saliency map explanations produced null or adverse effects on decision quality, particularly when explanations were imperfect or participants lacked the competence to assess them critically. Poursabzi-Sangdeh et al. (2021) show that explanations can create an illusion of understanding — users feel more confident without becoming more accurate — and may amplify cognitive biases by anchoring attention on highlighted features regardless of their diagnostic relevance. Von Zahn et al. (2024) further show that XAI's effects on metacognition depend strongly on cognitive engagement and explanation fidelity. These critiques impose important boundary conditions on the XAI-OLL framework: the sequential mechanism from XAI exposure to knowledge integration is not automatic. Corcoles-Jiménez et al. (2025) show that XAI disrupts traditional knowledge management, motivating the XAI-OLL framework.

2.2 Human-AI Collaboration and Trust Calibration

Effective human-AI collaboration requires that users neither over-rely on AI recommendations (automation bias) nor systematically reject them (algorithm aversion) (Dellermann et al., 2019). Lee and See (2004) define calibrated trust as a form of reliance proportionate to the system's actual reliability, updated dynamically through experience. Parasuraman et al. (1997) demonstrated that transparency-enhancing interventions reduce both over- and under-trust relative to opaque systems.

2.3 Knowledge Management, Organisational Learning, and GRAI Logics

Knowledge management scholarship has long investigated how organisations create, share, and apply knowledge across individual and collective levels (Nonaka and Takeuchi, 1995; Davenport and Prusak, 1998). The SECI model describes knowledge creation as a spiral involving four conversion modes: Socialisation (tacit-to-tacit), Externalisation (tacit-to-explicit), Combination (explicit-to-explicit), and Internalisation (explicit-to-tacit). AI systems are reshaping these dynamics. Jarrahi et al. (2023) argue that AI functions as an active knowledge partner, reshaping how professionals build expertise. Böhm and Durst (2025) propose that generative AI introduces new GRAI (Generative and Reflective AI Integration) logics of knowledge creation, extending the SECI spiral by positioning algorithmically generated knowledge as a distinct epistemic contribution.

2.4 Reskilling, Upskilling, and HR Transformation

The proliferation of AI in professional contexts is generating intense interest in workforce transformation. Morandini et al. (2023) identify the need for adaptive competencies in AI-augmented workplaces; Santana and Díaz-Fernández (2023) emphasise that technological progress demands higher-order capabilities such as critical judgment and interpretive competence. If XAI explanations systematically develop these competences, they represent an underutilised resource for embedded upskilling — delivered through the daily workflow rather than separate training. This reframes the HR professional's role from passive recipient to active designer of the conditions under which XAI-enabled learning can occur.

2.5 Reflective Practice and Dual-Process Reasoning

The cognitive mechanism proposed in XAI-OLL integrates two complementary traditions. Schön's (1983) theory of reflective practice posits that professional learning occurs through iterative cycles of action, observation, and reflection-on-action: moments in which practitioners consciously examine their own reasoning and assumptions in the light of new evidence. Kahneman's (2011) dual-process theory distinguishes between fast, heuristic cognition (System 1) and slow, deliberative reasoning (System 2). XAI explanations thus serve as an explicit analytical intermediary that may shift users from System 1 toward System 2, Together, these frameworks provide the micro-level cognitive foundation for the XAI-OLL sequential mechanism, augmented by von Zahn et al.'s (2024) demonstration that XAI exposure modulates metacognitive monitoring — an effect conditioned by explanation quality and the user's prior capacity for critical engagement.

3 The XAI-OLL Framework

3.1 Framework Overview

XAI-OLL (Explainable Artificial Intelligence as an Organisational Learning Lever) proposes that AI-generated explanations function not merely as transparency mechanisms but as cognitive and organisational levers that activate professional learning and knowledge integration. The framework positions XAI explainability as the antecedent of a sequential process comprising four constructs: Reflective Reasoning (RR), Interpretive Competence (IC), Calibrated Trust (CT), and Organisational Knowledge Integration (OKI). The model and its hypotheses are illustrated in Figure 1.

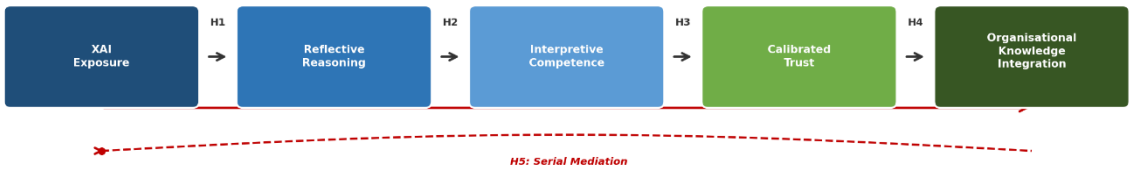


Figure 1. The XAI-OLL Sequential Model. Boxes represent latent constructs; labelled arrows indicate hypothesised causal paths (H1–H4); the dashed arc represents the serial mediation hypothesis (H5).

The framework articulates a specific relationship with the SECI model (Nonaka and Takeuchi, 1995) that requires careful conceptual specification. XAI explanations do not directly externalise the tacit knowledge of the human professional; rather, they make visible the tacit criteria embedded in the model's learned representations — a form of algorithmic externalisation. This visualisation of algorithmic logic then acts as a stimulus: confronted with an explicit representation of the model's reasoning, professionals are prompted to surface and articulate their own tacit interpretive criteria — the genuine Externalisation move in SECI terms. Through collaborative discussion of those explanations, teams subsequently synthesise individual explicit interpretations into shared diagnostic frameworks (Combination). The XAI-OLL framework thus provides the cognitive engine that GRAI logics (Böhm and Durst, 2025) describe at a structural level. Figure 2 illustrates the mapping between XAI-OLL and the SECI phases.

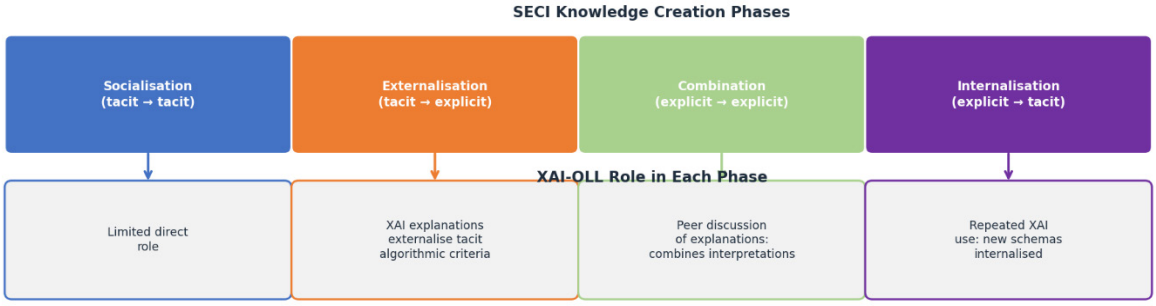


Figure 2. XAI-OLL within the SECI Knowledge Creation Framework. Upper row: SECI phases (Nonaka and Takeuchi, 1995); lower row: the corresponding role of XAI explanations in each phase.

3.2 Constructs and Formal Hypotheses

Reflective Reasoning (RR) is defined as the degree to which users engage in deliberate, comparative reasoning when examining XAI outputs — explicitly contrasting the model's rationale with their own prior assessment. The presence of a visual explanation disrupts heuristic acceptance (System 1) and creates an explicit cognitive reference point that invites comparison (System 2), enacting Schön's (1983) reflection-in-action:

H1: Exposure to XAI explanations positively activates reflective reasoning compared to prediction-only conditions.

Interpretive Competence (IC) is conceptually distinct from RR: whereas RR is an in-the-moment cognitive process (the deliberate comparison triggered by a specific explanation), IC is an acquired transferable skill — the durable capacity to critically engage with any AI-generated evidence across contexts, identifying relevant features, recognising model limitations, and discovering previously unnoticed patterns. IC accumulates through repeated reflective episodes and, once developed, persists beyond the immediate interaction — consistent with Ericsson et al.'s (1993) account of skill acquisition through deliberate practice and with Cheng et al.'s (2019) finding that XAI-based training transfers to new, unseen cases. Von Zahn et al. (2024) further demonstrate that XAI exposure stimulates metacognitive monitoring, supporting IC development — though conditioned by explanation fidelity. This aligns with the SECI Externalisation phase, where articulating one's tacit criteria in response to the algorithmic stimulus constitutes a genuine knowledge conversion event:

H2: Higher reflective reasoning positively predicts the development of interpretive competence.

Calibrated Trust (CT) represents proportionate reliance grounded in evidence quality rather than automation deference. Users with higher IC are better positioned to assess when AI recommendations are well-founded (Lee and See, 2004; Parasuraman et al., 1997). However, this relationship is conditional: Humer et al. (2024) show that imperfect explanations can attenuate or reverse the IC→CT effect by anchoring users on non-diagnostic features. XAI quality is therefore treated as a boundary condition moderating H3:

H3: Higher interpretive competence positively predicts calibrated trust (moderated by XAI explanation quality).

Organisational Knowledge Integration (OKI) captures the degree to which AI-derived insights contribute to shared professional knowledge through peer discussions, formalisation of criteria, and incorporation into practice. In this pilot, OKI is measured at team level (dyadic interaction), providing a proximal behavioural indicator; upward inference to the organisational level requires longitudinal team-based designs. OKI aligns with the Combination and Internalisation phases of SECI and with GRAI logics (Böhm and Durst, 2025):

H4: Higher calibrated trust positively predicts organisational knowledge integration.

H5: The effect of XAI explanations on organisational knowledge integration is serially mediated by reflective reasoning, interpretive competence, and calibrated trust.

In the pilot, H5 is treated as exploratory; confirmatory testing requires the full-scale study.

The path model implied by H1–H5 can be formally specified as a system of structural equations. Let $XAI_i \in \{0,1\}$ denote condition assignment; $Q_i \in [0,1]$ denote explanation quality; and RR_i, IC_i, CT_i, OKI_i denote the latent construct scores for participant i. The structural equations are: $RR_i = \beta_1 \cdot XAI_i + \zeta_1{}_i$ (H1); $IC_i = \beta_2 \cdot RR_i + \zeta_2{}_i$ (H2); $CT_i = \beta_3 \cdot IC_i + \beta_4 \cdot (IC_i \times Q_i) + \zeta_3{}_i$ (H3); $OKI_i = \beta_5 \cdot CT_i + \zeta_4{}_i$ (H4). The total indirect effect for H5 is $\delta = \beta_1\beta_2\beta_3\beta_5$, estimated via bootstrapped mediation analysis.

Table 1 summarises the five hypotheses alongside their theoretical bases. H3 is specified as a conditional relationship, with XAI explanation quality as an explicit moderator.

Table 1. XAI-OLL Hypotheses: Constructs, Theoretical Bases, and Expected Directions.

Hyp.	Relationship	Theoretical Basis	Direction
H1	XAI exposure → Reflective Reasoning	Kahneman (2011) dual-process; Schön (1983) reflective practice	Positive
H2	Reflective Reasoning → Interpretive Competence	Schön (1983); Hoffman et al. (2018); von Zahn et al. (2024); SECI Externalisation	Positive
H3	Interpretive Competence → Calibrated Trust (moderated by XAI quality)	Lee & See (2004); Parasuraman et al. (1997); Humer et al. (2024); Cai et al. (2024)	Positive (conditional)
H4	Calibrated Trust → Organisational Knowledge Integration	Nonaka & Takeuchi (1995); Jarrahi et al. (2023); GRAI logics	Positive
H5	XAI → OKI (serial mediation via RR, IC, CT)	Dual-process theory; XAI-OLL sequential mechanism	Positive indirect

4 Methodology

4.1 Research Design

This paper presents a theoretical-methodological contribution: the XAI-OLL framework and an associated pilot quasi-experimental protocol designed to operationalise its constructs and assess whether the hypothesised relationships are detectable before proceeding to full-scale validation. The pilot study is not yet executed; it constitutes the next planned step in this research programme. Participants will be randomly assigned to one of two between-subjects conditions:

Condition A (AI only): participants receive the model's probabilistic prediction (malignant/benign classification) without any explanation.

Condition B (AI + XAI): participants receive both the prediction and a visual explanation — a saliency map and Grad-CAM activation overlay — highlighting the image regions that drove the classification.

This two-condition design isolates the effect of explainability on the XAI-OLL constructs while controlling for all other aspects of the AI interaction, and provides a direct test of whether XAI is a necessary ingredient or merely one of several possible reflective stimuli. Independent assessment before AI exposure (Phase 1) allows measurement of decision revision as a behavioural indicator of cognitive engagement. Explanation quality will be operationalised as a moderator of H3 by stratifying cases into high- and low-fidelity subconditions. Crucially, the multi-phase design allows distinguishing perceived competence (self-report) from actual performance improvement (accuracy on transfer cases), strengthening the construct validity of IC.

4.2 Context and Participants

The experimental context is the interpretation of histopathology slide images, a domain in which deep learning models achieve diagnostic accuracy comparable to specialist pathologists and in which XAI visualisation techniques are technically mature. The domain is characterised by high epistemic stakes, well-defined interpretive criteria, and professional communities accustomed to collaborative case review — making it an ideal setting for studying learning dynamics in human-AI interaction.

The pilot sample will consist of approximately N = 50 participants from medical and biomedical communities: pathologists, resident physicians, biomedical researchers, and advanced medical students. Variation in professional background allows exploration of expertise-level moderation effects (Parasuraman et al., 1997). This size is appropriate for PLS-SEM pilot analysis; a subsequent study will target N ≥ 200 for covariance-based SEM (Hair et al., 2019).

4.3 Procedure and Measures

The procedure will unfold in four sequential phases. In Phase 1 (Independent Assessment), participants observe histopathology images and record their diagnostic judgment and confidence on a seven-point scale, without AI assistance. In Phase 2 (AI Exposure), the system's output is presented according to condition; participants may revise their diagnosis. Behavioural data — interaction time, decision revision rate, and accuracy against ground-truth labels — are logged automatically, providing objective performance indicators alongside self-report. Calibration error (the difference between stated confidence and actual accuracy) is computed per participant as an objective CT proxy, following Bansal et al. (2021). In Phase 3 (Reflective Task), participants provide a written justification and complete three unseen transfer cases without AI assistance — providing a behavioural IC measure (actual accuracy improvement) that complements Likert self-report and allows distinguishing perceived from actual competence. In Phase 4 (OKI Group Task), dyads discuss one shared case

and reach a joint diagnosis, providing a behavioural proxy for team-level knowledge integration beyond individual perception.

The four XAI-OLL constructs are operationalised as seven-point Likert scales. Table 2 presents sample items and sources.

Table 2. XAI-OLL Constructs: Sample Measurement Items and Theoretical Sources.

Construct	Sample Measurement Items (7-pt Likert)	Key Sources
Reflective Reasoning (RR)	"I consciously compared my interpretation with the AI-highlighted features"; "Viewing the explanation caused me to revisit my initial assessment"	Schön (1983); Kahneman (2011)
Interpretive Competence (IC)	"The explanation drew my attention to features I had not previously considered"; "I can now better identify diagnostically relevant image regions"	Hoffman et al. (2018); von Zahn et al. (2024); Arrieta et al. (2020)
Calibrated Trust (CT)	"I evaluated the AI's prediction critically rather than accepting it automatically"; "I could identify cases where the AI's reasoning seemed unreliable"	Lee & See (2004); Cai et al. (2024); Humer et al. (2024)
Organisational Knowledge Integration (OKI)	"The AI explanation would be useful to discuss with colleagues"; "This type of explanation could improve our team's shared diagnostic criteria"	Nonaka & Takeuchi (1995); Star & Griesemer (1989)

4.4 Analysis Strategy

Quantitative data will be analysed using PLS-SEM, well-suited to small pilot samples and exploratory model testing (Hair et al., 2019). Measurement quality will be assessed via Cronbach's alpha, composite reliability, AVE, and HTMT discriminant validity. The structural model will be evaluated through path coefficients and bootstrapped indirect effects (5,000 resamples) for H5. To address common method bias, self-reported constructs will be triangulated with objective behavioural indicators: accuracy improvement (Phase 1 vs. Phase 2 against ground-truth labels) as an IC proxy, and calibration error as a CT proxy. The moderation of H3 will be tested through subgroup analysis across high- and low-fidelity explanation conditions. Qualitative texts from Phase 3 and dyadic interaction data from Phase 4 will be coded using thematic analysis to contextualise quantitative findings.

5 Anticipated Propositions and Theoretical Discussion

5.1 Anticipated Empirical Patterns

On the basis of the theoretical model and existing literature, we formulate the following anticipated patterns to guide empirical interpretation. Regarding H1, participants in Condition B (XAI present) are expected to report significantly higher Reflective Reasoning scores than those in Condition A. Visual explanations disrupt heuristic acceptance and invite deliberate comparison — activating Kahneman's (2011) System 2 and enacting Schön's (1983) reflection-in-action. Regarding H2, IC is anticipated to emerge as users verbalise their diagnostic reasoning against XAI-highlighted features, externalising and refining interpretive schemas — enacting the SECI Externalisation mode, consistent with von Zahn et al. (2024).

Regarding H3, the IC→CT relationship is anticipated to be positive but conditional on explanation quality. Where Grad-CAM overlays faithfully reflect diagnostically relevant regions, users should demonstrate more accurate trust calibration, avoiding automation bias and algorithm aversion (Lee and See, 2004). Consistent with Humer et al. (2024), low-fidelity conditions are expected to attenuate or reverse this effect — a critical warning for organisations deploying XAI without quality assurance.

Regarding H4 and H5, Calibrated Trust is anticipated to predict OKI, and the serial mediation (H5) constitutes the XAI-OLL's most theoretically original proposition: that XAI affects organisational learning through a sequential cognitive chain, not a single mechanism — formalising conceptually what GRAI logics (Böhm and Durst, 2025) describe at a structural level.

5.2 XAI Explanations as Boundary Objects and Learning Levers

A key theoretical implication of the XAI-OLL framework concerns the social and collective dimensions of AI-generated explanations. Activation maps and local attribution outputs — particularly in visual domains such as histopathology — make visible what would otherwise remain tacit algorithmic knowledge, creating cognitive stimuli that prompt professionals to surface and articulate their own tacit interpretive criteria. In peer review

settings, these artefacts become objects of collective scrutiny: clinicians compare highlighted regions, challenge each other's interpretations, and arrive at shared diagnostic criteria — a Combination dynamic in the SECI spiral. Through repeated use, new interpretive schemas become internalised into professional practice (Internalisation). XAI explanations thus function as boundary objects (Star and Griesemer, 1989), bridging the epistemic gap between algorithmic and human knowledge communities, enabling the knowledge co-construction that GRAI logics theorise at the structural level.

6 Organisational Conditions for XAI-Enabled Upskilling

The XAI-OLL framework does not predict that XAI explanations will automatically generate organisational learning. The evidence reviewed — particularly Humer et al. (2024) and von Zahn et al. (2024) — indicates that the pathway from XAI exposure to knowledge integration depends on enabling conditions that organisations must actively design.

Designing for interpretive competence development. Professional training that incorporates XAI explanations as case-based learning materials is a prerequisite for activating the XAI-OLL mechanism. HR professionals must incorporate interpretive competence — not merely operational fluency with AI interfaces — as a core AI literacy dimension (Morandini et al., 2023; Santana and Díaz-Fernández, 2023), redesigning development pathways to include structured practice with XAI tools.

Governance and quality assurance of XAI tools. Because H3 is conditioned by explanation quality, organisations must establish governance standards: protocols for vetting explanation fidelity, regular audits against domain benchmarks, and counter-anchoring training that encourages independent assessment before AI exposure.

Creating knowledge integration infrastructure. OKI requires formalised channels for moving individual AI-assisted insights into collective knowledge: regular review sessions, structured processes for updating practice guidelines, and communities of practice in which professionals share interpretive experiences with AI tools. This 'learning in the flow of work' dynamic (Jarrahi et al., 2023) is organisationally designed, not spontaneous.

7 Theoretical Contributions and Limitations

This paper makes three theoretical contributions. First, it presents the first formal sequential cognitive-organisational model of how XAI explainability generates professional learning and organisational knowledge integration, specifying a testable sequential pathway (H1–H5) and incorporating boundary conditions on H3, engaging with the contradictory empirical literature (Humer et al., 2024; von Zahn et al., 2024). Second, the paper extends the SECI model with a precise specification: XAI functions as an algorithmic externalisation stimulus that provokes professional Externalisation rather than constituting it directly — clarifying a conceptual ambiguity in prior AI-knowledge management research. Third, the paper situates XAI within the HR transformation literature, positioning AI interpretive literacy as a new professional competence dimension, and HR professionals as active designers of AI-enabled upskilling conditions (Morandini et al., 2023).

The framework is currently theoretical: no empirical data have yet been collected. The pilot, once executed, will be constrained by small sample size and single-domain focus (histopathology), limiting causal inference and generalisability. Future research should pursue longitudinal designs, larger samples ($N \geq 200$), and cross-domain comparative studies.

8 Conclusions

The XAI-OLL framework proposes a theoretically grounded and empirically operationalisable account of how XAI explanations function as organisational learning levers rather than mere transparency tools. By specifying a sequential cognitive-organisational mechanism — XAI Exposure → Reflective Reasoning → Interpretive Competence → Calibrated Trust → Organisational Knowledge Integration — the framework generates five testable propositions grounded in knowledge management theory and operationalised through a formal SEM specification and a pilot quasi-experimental protocol. The framework makes three interconnected contributions: it advances mechanism-level theory of XAI-driven learning; it extends the SECI model into human-AI collaboration with conceptual precision; and it repositions HR professionals as active architects of AI-enabled upskilling environments. Critically, the framework does not assume that XAI automatically generates learning — it identifies the boundary conditions under which the pathway holds and the organisational investments required to activate it. As AI decision support becomes pervasive across professional domains, understanding how organisations can harness explanatory outputs as learning infrastructure represents a critical frontier for knowledge management theory and HR practice alike.

9 Ethics Declaration

The planned empirical study involves human participants. Informed consent will be obtained from all participants prior to data collection. No clinical patient data will be used; images will be drawn exclusively from publicly available, anonymised histopathology datasets. Participation will be voluntary, anonymous, and may be withdrawn at any time without consequence. Ethical approval will be sought from the relevant institutional review board prior to data collection.

10 AI Declaration

AI tools were used to support literature identification, drafting, and editing during the preparation of this paper. All theoretical constructs, framework design, hypotheses, and interpretations are the original work of the author. Full responsibility for the content rests with the author.

References

Arrieta, A.B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., García, S., Gil-López, S., Molina, D., Benjamins, R., Chatila, R. and Herrera, F. (2020) “Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI”, Information Fusion, Vol. 58, pp. 82–115.

Bansal, G., Nushi, B., Kamar, E., Weld, D.S., Lasecki, W.S. and Horvitz, E. (2021) "Does the Whole Exceed its Parts? The Effect of AI Explanations on Complementary Team Performance", Proceedings of the ACM CHI Conference on Human Factors in Computing Systems, Article 81.

Böhm, I. and Durst, S. (2025) “Revisiting the SECI Model: The Impact of Generative AI on Organisational Knowledge Creation”, in Proceedings of IFKAD 2025, Budapest.

Cai, C.J., Jongejan, J. and Holbrook, J. (2024) “C-XAI: A Conceptual Framework for Designing XAI Tools that Support Trust Calibration”, Computers in Human Behavior Reports, Vol. 13.

Cheng, H.F., Wang, R., Zhang, Z., O'Connell, F., Gray, T., Harper, F.M. and Zhu, H. (2019) "Explaining Decision-Making Algorithms through UI: Strategies to Help Non-Expert Stakeholders", Proceedings of the ACM CHI Conference on Human Factors in Computing Systems, Article 559.

Chromik, M. and Schubert, T. (2025) “Less than one percent: The human evaluation gap in explainable AI”, arXiv preprint arXiv:2503.16507.

Corcoles-Jiménez, M., Villena-Martínez, V., Bueno-Crespo, A. and Morales-García, J. (2025) “The role of explainable artificial intelligence (XAI) in innovation processes: a knowledge management perspective”, Technology in Society, Vol. 81.

Davenport, T.H. and Prusak, L. (1998) Working Knowledge: How Organisations Manage What They Know, Harvard Business School Press, Boston.

Dellermann, D., Ebel, P., Söllner, M. and Leimeister, J.M. (2019) “Hybrid Intelligence”, Business and Information Systems Engineering, Vol. 61, No. 5, pp. 637–643.

Doshi-Velez, F. and Kim, B. (2017) “Towards a rigorous science of interpretable machine learning”, arXiv preprint arXiv:1702.08608.

Ericsson, K.A., Krampe, R.T. and Tesch-Romer, C. (1993) "The Role of Deliberate Practice in the Acquisition of Expert Performance", Psychological Review, Vol. 100, No. 3, pp. 363-406.

Hair, J.F., Risher, J.J., Sarstedt, M. and Ringle, C.M. (2019) “When to use and how to report the results of PLS-SEM”, European Business Review, Vol. 31, No. 1, pp. 2–24.

Hoffman, R.R., Mueller, S.T., Klein, G. and Litman, J. (2018) “Metrics for Explainable AI: Challenges and Prospects”, arXiv preprint arXiv:1812.04608.

Humer, C., Heitz, C., Jain, N. and Feuerriegel, S. (2024) “Do saliency maps improve understanding? Adverse effects of XAI explanations on decision quality in a human-in-the-loop context”, Proceedings of the ACM CHI Conference on Human Factors in Computing Systems.

Jarrahi, M.H., Askay, D., Eshraghi, A. and Smith, P. (2023) “Artificial intelligence and knowledge management: A partnership between human and AI”, Business Horizons, Vol. 66, No. 1, pp. 87–99.

Kahneman, D. (2011) Thinking, Fast and Slow, Farrar, Straus and Giroux, New York.

Lee, J.D. and See, K.A. (2004) “Trust in automation: Designing for appropriate reliance”, Human Factors, Vol. 46, No. 1, pp. 50–80.

Longo, L., Brcic, M., Cabitza, F., Choi, J., Confalonieri, R., Del Ser, J., Guidotti, R., Hayashi, Y., Herrera, F., Holzinger, A. and Jiang, R. (2024) “Explainable Artificial Intelligence (XAI) 2.0: A manifesto of open challenges and interdisciplinary research directions”, Information Fusion, Vol. 106.

Lundberg, S.M. and Lee, S.I. (2017) “A unified approach to interpreting model predictions”, Advances in Neural Information Processing Systems, Vol. 30, pp. 4765–4774.

Morandini, S., Fraboni, F., De Angelis, M., Nardelli, M., Di Nicola, L. and Giusino, D. (2023) “The Impact of Artificial Intelligence on Workers’ Skills: Upskilling and Reskilling in Organisations”, Informatics, Vol. 10, No. 2.

Nonaka, I. and Takeuchi, H. (1995) The Knowledge-Creating Company, Oxford University Press, New York.

Parasuraman, R., Sheridan, T.B. and Wickens, C.D. (1997) “A model for types and levels of human interaction with automation”, IEEE Transactions on Systems, Man, and Cybernetics, Vol. 30, No. 3, pp. 286–297.

Poursabzi-Sangdeh, F., Goldstein, D.G., Hofman, J.M., Vaughan, J.W. and Wallach, H. (2021) “Manipulating and Measuring Model Interpretability”, Proceedings of the ACM CHI Conference on Human Factors in Computing Systems, Article 237.

Ribeiro, M.T., Singh, S. and Guestrin, C. (2016) “’Why should I trust you?’: Explaining the predictions of any classifier”, Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 1135–1144.”

Santana, M. and Díaz-Fernández, M. (2023) “State of the art and future directions of the competencies for the new landscape of AI-augmented HRM”, International Journal of Human Resource Management, Vol. 34, No. 11, pp. 2138–2166.

Schön, D.A. (1983) The Reflective Practitioner: How Professionals Think in Action, Basic Books, New York.

Selvaraju, R.R., Cogswell, M., Das, A., Vedantam, R., Parikh, D. and Batra, D. (2017) “Grad-CAM: Visual explanations from deep networks via gradient-based localization”, Proceedings of the IEEE International Conference on Computer Vision, pp. 618–626.

Star, S.L. and Griesemer, J.R. (1989) “Institutional ecology, ‘translations’ and boundary objects: Amateurs and professionals in Berkeley’s Museum of Vertebrate Zoology”, Social Studies of Science, Vol. 19, No. 3, pp. 387–420.

Von Krogh, G. (2018) “Artificial intelligence in organizations: New opportunities for phenomenon-based theorizing”, Academy of Management Discoveries, Vol. 4, No. 4, pp. 404–409.

Von Zahn, M., Feuerriegel, S. and Kühl, N. (2024) “The Business Value of Explainable AI: An Experimental Study of XAI and Metacognition”, Information Systems Research, Vol. 35, No. 2, pp. 673–696.

Tacit Knowledge Loss (TKL) from an Intelligent Knowledge-Based View (IKBV)

Joachim Lorenz
Reykjavik University, Reykjavik, Iceland

Abstract

Tacit knowledge loss (TKL) poses a critical challenge to an organization's competitiveness, especially when experienced experts retire. This challenge is exacerbated by the high cost of regenerating lost knowledge within organizations. Using an Intelligent Knowledge-Based View (IKBV) framework that integrates human and artificial intelligence, this paper explores how strategic approaches to knowledge retention, supported by generative AI (GenAI), can effectively mitigate TKL. A scoping review was conducted for the study using Scopus as the database. A total of 5 studies were identified to address the three research questions. This study suggests that strategic approaches to knowledge retention, such as capture, storage, and retrieval mechanisms, can be combined with GenAI methods to capture and synthesize the knowledge of mobile experts (RQ1). The resulting AI-supported knowledge retention mechanisms can be strategically complemented with an intelligence-based perspective (RQ2) to identify the critical types of knowledge and the mobile experts who need to be involved in knowledge retention. The IKBV framework resulting from the study (RQ3) can be understood as an intelligent knowledge retention strategy framework that incorporates all both research questions (RQ1; RQ2). The paper expands the theoretical understanding of knowledge loss research by proposing that future research investigate intelligent knowledge retention mechanisms to further test their effectiveness in mitigating knowledge loss. Furthermore, the study revealed that the mobile term "expert" needs to be more precisely defined so that researchers can specify the target group in their empirical studies. This is where a doctoral thesis could begin exploring how the concept of an expert can be defined and differentiated within an organization, enabling sustainable prioritization within an intelligent knowledge retention strategy. The findings show that a knowledge retention strategy requires not only technological implementation, but also human expertise and intelligent technological interaction. The knowledge retention strategy should be prioritized in organizations where the loss of intellectual capital due to the departure of experts can become a risk to innovation. This paper contributes to the theory of knowledge management by showing how the IKBV framework can bridge the gap between human expert knowledge and artificial intelligence capabilities, thus promoting sustainable organizational learning.

Keywords: Tacit Knowledge Loss, Knowledge Retention, Intelligent Knowledge-Based View, Generative Artificial Intelligence, Mobile Experts

Paper type: Academic Research Paper

1 Introduction

The loss of tacit knowledge is one of the most significant, but underestimated, organizational risks in the modern knowledge economy.

As Polanyi (1966) put it: “we know more than we can tell” (p. 4), capturing the essence of why tacit knowledge is so difficult to transmit, document, and preserve. This fundamental challenge becomes crucial when experienced employees leave an organization, taking with them years of accumulated experience, contextual understanding, and implicit expertise that cannot be easily replicated through documentation or training programs.

1.1 Problem statement

The literature on knowledge management has long recognized that organizational competitive advantage increasingly comes from intangible assets, especially human capital and organizational knowledge (Eisenhardt and Santos, 2012; Nonaka and von Krogh, 2009).

In this context, the tacit loss of knowledge (TKL) – the departure of critical knowledge when experts leave the organization – threatens organizational performance and innovation. Research by Massingham (2018) shows that the regeneration of lost tacit knowledge requires significant financial and time investments, as new knowledge must be created through costly interactions with social actors and organizational learning processes (Nonaka and Takeuchi, 1995).

Recent developments in generative artificial intelligence (GenAI) have created unprecedented opportunities for practice in knowledge management. Alavi et al. (2024) argue that GenAI technologies enable novel approaches to capture, synthesize, and mobilize organizational knowledge.

Despite this technological promise, a theoretical problem is that there is limited research on how strategic approaches can effectively leverage intelligent systems to mitigate TKL, especially in the context of departing experts.

1.2 Purpose statement

This paper approaches this research problem by presenting intelligent mechanisms to address TKL. The resulting Intelligent Knowledge-Based View (IKBV) is primarily intended to help researchers examine AI-supported knowledge retention mechanisms so that organizations can effectively counter the threat of knowledge loss. The IKBV expands the traditional knowledge-based view (KBV) by recognizing that intelligent knowledge retention requires not only human-centered or technology-centered approaches, but rather hybrid systems that leverage the strengths of both (Kaczorowska-Spychalska *et al.*, 2024; Lichtenthaler, 2019).

Furthermore, the IKBV includes the Knowledge Retention Cycle by Levallet and Chan (2019) and expands it with an intelligent, knowledge-based perspective to counter the problems identified that increase the risk of knowledge loss. The focus of the consideration should be on individual experiential knowledge, i.e. the knowledge that the organization needs to maintain its innovative strength.

The author of this study, following Levallet and Chan (2019), defines this knowledge as mobile expert knowledge and loss of knowledge as a result when this mobile expert knowledge is lost due to the departure of mobile experts.

Consequently, the central research question (RQ) is: Which intelligent knowledge retention mechanisms can counteract the loss of mobile expert knowledge?

The following sub-questions are to be answered:

- *RQ1: What role does GenAI play within a knowledge retention strategy?*
- *RQ2: Which model can be employed for the expansion of the Knowledge-Based View (KBV)?*
- *RQ3: How can an intelligent knowledge retention strategy framework be designed that incorporates all two research questions?*

2 Background

2.1 The knowledge-based view

The Knowledge-Based View (KBV) of strategic management states that organizational knowledge and skills are the primary sources of competitive advantage (Eisenhardt and Santos, 2012). Within the KBV framework, knowledge retention (KR) emerges as a critical strategic capability – the ability to retain and use existing tacit knowledge from departing experts to maintain a competitive position and support organizational continuity.

2.2 Tacit knowledge and organizational value

Tacit knowledge—the experiential, context-dependent understanding that individuals develop through practice and social interaction—makes up a significant portion of organizational knowledge. Unlike explicit knowledge, which can be codified in documents, databases, or manuals, tacit knowledge remains embedded in human cognitive and behavioral patterns. Nonaka and Takeuchi (1995) point out that tacit knowledge encompasses mental models, beliefs, perspectives, and expertise developed through experience and practice.

The strategic importance of tacit knowledge stems from its difficulty of imitation and its central importance for competitive advantage (Eisenhardt and Santos, 2012; Zack, 1999). Experienced engineers, managers, craftsmen, and long-term professionals accumulate tacit knowledge that shapes organizational practices, decision-making processes, and innovation skills.

When experts retire, organizations face the immediate challenge of knowledge loss and the long-term costs of reconstructing that knowledge. Massingham (2008) empirically shows that knowledge loss affects organizational performance and innovation. Research shows that regenerating lost tacit knowledge requires significant organizational investment, as knowledge creation is inherently a social process that involves interaction, dialogue, and learning.

This finding underscores the strategic need for proactive knowledge retention strategies that capture expert knowledge before departure.

2.3 Knowledge retention cycle

The core consideration of the present study is based on the Revised Knowledge Retention (KR) Cycle by Levallet and Chan (2019). The original KR Cycle (p. 179) includes three phases: Knowledge Capture, Knowledge Storage, and Knowledge Retrieval. The loss of knowledge is presented here as a phenomenon that results from an inability to record and store the individual knowledge of mobile experts for the organization.

The extended KR Cycle (p. 192) distinguishes between individual and social knowledge. The model can be used to explain how the individual knowledge of a mobile, or more precisely, retired expert, can be transformed into an organizational asset. The authors argue for knowledge transfer mechanisms that enable transformation into social knowledge, which should also be recorded, stored, and made available.

Here, an inability to transform implicit individual knowledge (automatic knowledge) into social knowledge can increase the risk of knowledge loss. This cyclical model provides a framework for understanding how organizations can systematically address knowledge retention and the risks of not having this capability. One limitation of the model is that it does not show how emerging technologies, especially generative AI, can support each stage of the retention cycle.

3 Methodology

This research uses a scoping review methodology to synthesize existing literature on strategic approaches to tacit knowledge loss and the role of intelligent knowledge systems. A scoping review provides a systematic approach to identifying, analyzing, and synthesizing evidence across a broad thematic area (Westphaln *et al.*, 2021).

To answer the research questions, Scopus was selected as the database, and the following search string was used as the search strategy:

TITLE-ABS-KEY(("generative artificial intelligence" OR "generative AI" OR "GenAI")
AND ("knowledge management" OR "knowledge retention" OR "tacit knowledge"))

This identified 547 studies. For the sake of consolidation and due to the limited scope of this conference paper, the corpus has been reduced: business, management and accounting (n = 100); article (n = 192); English (n = 540); journal (n = 217); Journal of Knowledge Management (n = 9) and VINE Journal of Information and Knowledge Management System (n = 3). Only one source could be included: Böhm and Durst (2025). The backward citation searching has yielded four additional sources; see Table 1.

Table 1**:** Included Studies

Scoping Review & Backward Citation Searching		
Author	Year	Title
Alavi et al.	2024	A Knowledge Management Perspective of Generative Artificial Intelligence
Böhm and Durst	2025	Knowledge management in the age of generative artificial intelligence – from SECI to GRAI
Ismail Al-Alawi and Al-Ahmed	2025	Integration of AI in Capturing Tacit Knowledge of Employees Leading to Innovation in Organizational Learning: A Literature Review
Rezaei	2025	Artificial intelligence in knowledge management: Identifying and addressing the key implementation challenges
Lichtenthaler	2019	An Intelligence-Based View of Firm Performance: Profiting from Artificial Intelligence

4 Findings

4.1 RQ1: What role does GenAI play within a knowledge retention strategy?

Generative artificial intelligence represents a technological paradigm shift in knowledge management. Alavi et al. (2024) argue that GenAI enables novel approaches to knowledge synthesis, availability, and application. Unlike traditional knowledge management systems, which mainly store and retrieve existing knowledge, GenAI systems can generate new synthetic knowledge by combining multiple sources of information and applying sophisticated pattern recognition.

Böhm and Durst (2025) propose a model that evolves from the Nonaka-Takeuchi SECI (Socialization, Externalization, Combination, Internalization) framework to a GRAI framework that integrates generative AI thinking and intelligence. This development recognizes that GenAI can facilitate knowledge externalization by

helping experts articulate tacit knowledge through dialogue, enhance combination by synthesizing dispersed knowledge sources, and support internalization through personalized learning experiences.

Ismail Al-Alawi and Al-Ahmed (2025) empirically demonstrate that AI-powered systems can improve organizational decision-making by leveraging predictive analytics to capture and apply employee knowledge. Her research highlights that transparent methods for capturing and synthesizing knowledge – instead of "black-box" AI implementations – create organizational conditions for sustainable knowledge cultures.

Rezaei (2025) shows how knowledge storage (KST) can be understood conceptually. Thus, KST is understood "the intricate processes and technological infrastructure employed for capturing, preserving, organising, and retrieving knowledge within an organisation" (p. 3). The KST aims to ensure that knowledge is easily accessible and usable. According to the author, KST primarily refers to an IT-supported form, which can be implemented using a data warehouse, for example. This involves collecting large amounts of data, cataloging them, and generating knowledge.

4.1.1 Intelligent knowledge capture

GenAI systems can facilitate knowledge capture by involving experts in a dialogue that increasingly externalizes tacit knowledge. Large language models can conduct interviews, ask clarifying questions, and help experts formulate an implicit understanding.

Alavi et al. (2024) discuss the potential of GenAI to enable the organization to generate new knowledge from external sources or by automatically capturing meeting notes.

Böhm and Durst (2025) point out that this dialogue-based externalization overcomes the limitations of traditional documentation approaches by creating dynamic, contextual representations of knowledge. The authors discuss the potential of GenAI as semantic intelligence that can support the organization in capturing tacit knowledge, for example, by linking data or identifying relationships.

Rezaei (2025) also understands this possibility of capturing new knowledge by accessing various sources.

Ismail Al-Alawi and Al-Ahmed (2025) postulate the existence of knowledge management systems capable of capturing tacit knowledge.

4.1.2 Intelligent knowledge storage

GenAI enables combining knowledge from multiple experts, disparate documents, and disparate organizational sources to create synthetic knowledge artifacts that transcend any single source. Alavi et al. (2024) emphasize that GenAI can synthesize patterns, best practices, and integrated insights from dispersed knowledge sources, thus creating comprehensive organizational knowledge that represents collective rather than individual expertise. Rezaei (2025) discusses, among other things, algorithms that can clean the database, freeing up more storage space.

4.1.3 Intelligent knowledge retrieval

GenAI systems can support the knowledge queries and applications via conversational interfaces, recommender systems, and decision support tools. These technologies can help organizational members quickly acquire relevant knowledge and receive contextual guidance for specific situations and decisions.

Böhm and Durst (2025) discuss the retrieval-augmented generation (RAG) functionality, which enables AI to access an externalized document collection. This makes it easier to call up information. Ismail Al-Alawi and Al-Ahmed (2025) show that AI-powered predictive analytics can facilitate proactive knowledge application in decision-making contexts. According to Rezaei (2025), search experiences can also be personalized. In this case, the AI model learns the user's preferences and behavior patterns.

4.2 RQ2: Which model can be employed for the expansion of the Knowledge-Based View (KBV)?

Lichtenthaler (2019) suggests that contemporary organizations need an 'intelligence-based view' that integrates perspectives from multiple forms of intelligence – human, organizational, and artificial. This framework recognizes that modern competitive advantage increasingly stems from the ability to combine human expertise with artificial intelligence in complementary ways.

Building on this conceptual foundation, the Intelligent Knowledge-Based View (IKBV) framework extends the traditional P/B ratio by explicitly integrating hybrid human-AI knowledge systems into the theory of competitive advantage. The IKBV assumes that sustainable competitive advantage increasingly results from the ability of organizations to:

Identify, codify, and retain the tacit knowledge of experts through human-centered knowledge gathering processes.

Aggregate captured knowledge into organizational assets through GenAI-assisted analysis and synthesis.

Deploy hybrid systems that mobilize both human judgment and AI-powered decision support to apply knowledge in organizational contexts.

5 RQ3: How can an intelligent knowledge retention strategy framework be designed that incorporates both research questions?

The research results suggest that AI-supported KR mechanisms have already been established alongside traditional knowledge management methods. With the help of the scoping review, three technological approaches were identified for the three primary approaches to KR, according to (Levallet and Chan, 2019): knowledge capture, knowledge storage, and knowledge retrieval: semantic intelligence, algorithm, and RAG. While semantic intelligence can combine existing and external knowledge, algorithms aim to prepare this knowledge so it can be retrieved via a RAG.

However, this represents a purely technical view of knowledge retention and requires human control mechanisms, as GenAI tends to be biased by data (Alavi *et al.*, 2024). There is also the question of the extent to which these AI-supported mechanisms can identify mobile expert knowledge when it is scattered or not yet codified. Consequently, human mechanisms and thus classical methods are also needed as part of a knowledge retention strategy, so that mobile expert knowledge can be recorded and retained through interviews (Ismail Al-Alawi and Al-Ahmed, 2025).

Finally, in addition to intelligent KR mechanisms and classical methods, there is also a view of which types of knowledge should be retained. This study cannot answer this question. Also, the question of how an expert can be defined or even identified. The IKBV must therefore be expanded to include these issues. In the figure below, the author proposes an IKBV framework.

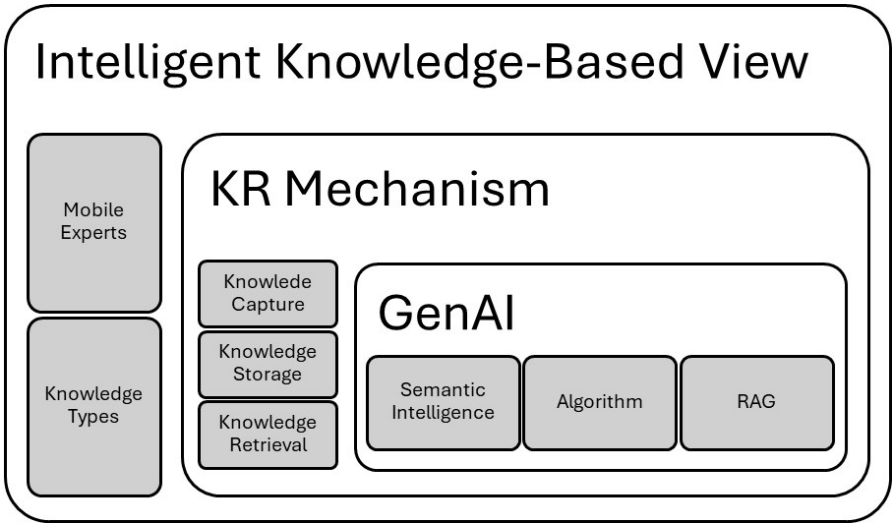


Figure 1: The Intelligent Knowledge-Based View (IKBV)

6 Implications

6.1 Implications for knowledge management theory

The IKBV framework contributes to knowledge management theory by extending the traditional knowledge-based view to include artificial intelligence as a strategic organizational capability. Rather than treating technology as a peripheral support function, the framework recognizes that GenAI-powered systems are increasingly central strategic assets for knowledge management in the modern organization.

This framework integrates GenAI into a coherent strategic approach to KR. Instead of viewing GenAI as a replacement for human knowledge management, the IKBV positions itself as a complementary capability that enhances organizational KR throughout the entire KR cycle. The results also support the development of human-centered SECI knowledge creation models towards integrated human-AI knowledge models, as proposed by Böhm and Durst (2025).

6.2 Practical implications for organizations

Organizations that want to mitigate the tacit loss of knowledge should adopt integrated approaches that combine human-centered and AI-powered strategies. Specifically, effective knowledge retention strategies should:

Knowledge capture: The use of semantic intelligence to unlock and link different knowledge bases so that mobile expert knowledge can be intelligently captured.

Knowledge Storage: The use of algorithms so that this knowledge base can be organized and intelligently structured.

Knowledge retrieval: The use of a RAG so that knowledge can be retrieved intelligently.

7 Conclusion and Future Research Directions

The loss of tacit knowledge is a critical organizational challenge in knowledge-intensive industries and sectors. The mobile expert phenomenon threatens organizational knowledge, competitive advantage, and innovation potential. Traditional approaches to knowledge storage partially address this challenge, but struggle with the inherently tacit, contextual, and embodied nature of mobile expert knowledge.

In this study, the research interest was in identifying which intelligent knowledge retention mechanisms can counteract the loss of mobile expert knowledge. Due to data scarcity, this question can only be answered to a limited extent. The study proposes to include two models in this consideration. On the one hand, the KR Cycle by Levallet and Chan (2019) offers a framework for differentiating KR into three parts: knowledge capture, knowledge storage, and knowledge retrieval. For the knowledge-based lens, Lichtenthaler's (2019) IBV can be used to expand the KBV.

The IKBV developed as part of this study uses these two models and gives an impression of how an intelligent view of knowledge retention can be designed. As organizations increasingly compete for knowledge and innovation, the ability to retain mobile expert knowledge is proving to be a critical strategic skill.

The IKBV framework provides a theoretical foundation and practical guidance for organizations seeking to protect knowledge assets, support organizational continuity, and secure a competitive advantage amid inevitable expert departures. Future research that validates this framework in different organizational contexts will further consolidate the usefulness of IKBV for the practice and theory of knowledge management.

Mixed-methods designs could be used in future research designs to assess the effectiveness of IKBV-oriented interventions in reducing knowledge loss and supporting organizational performance. Second, important questions remain about data management, knowledge ownership, and the ethical implications of using GenAI for knowledge capture and synthesis. Especially in contexts where the knowledge of mobile experts is captured, organizations need to address issues of consent, intellectual property, and fair use of synthetic knowledge resources. Third, the integration of IKBV approaches must go beyond sharing expert knowledge and address broader organizational learning, innovation, and knowledge sharing.

Future research should explore how IKBV frameworks support organizational capabilities beyond knowledge retention, such as in identifying critical mobile expertise.

8 Acknowledgements

The author acknowledges the support from the University of Reykjavík's research infrastructure, as well as the valuable feedback from the author's supervisor, Dr. Susanne Durst, and colleagues in the knowledge management research community.

9 Ethics Declaration

This research was limited to a theoretical review of scientific literature and did not include human subjects, data collection from individuals, or sensitive personal information. Therefore, ethical approval by an institutional review panel was not required. The research adheres to established academic standards for literature review and synthesis.

10 AI Declaration

Claude (Haiku Sonnet 4.5) was used to critically reflect on the paper's structure. Grammarly was used to improve spelling and expression.

References

Alavi, M., Leidner, D.E. and Mousavi, R. (2024), "Knowledge Management Perspective of Generative Artificial Intelligence", *Journal of the Association for Information Systems*, doi: 10.17705/1jais.00859.

Böhm, K. and Durst, S. (2025), "Knowledge management in the age of generative artificial intelligence – from SECI to GRAI", *VINE Journal of Information and Knowledge Management Systems*, Emerald Publishing, pp. 2059–5891, doi: 10.1108/VJIKMS-10-2024-0357/FULL/PDF.

Eisenhardt, K.M. and Santos, F.M. (2012), "Knowledge-Based View: A New Theory of Strategy?", *Handbook of Strategy and Management*, doi: 10.4135/9781848608313.n7.

Ismail Al-Alawi, A. and Al-Ahmed, S.A. (2025), "Integration of AI in Capturing Tacit Knowledge of Employees Leading to Innovation in Organizational Learning: A Literature Review", *2nd International Conference on IT Innovations and Knowledge Discovery*, *ITIKD 2024*, doi: 10.1109/ITIKD63574.2025.11005249.

Kaczorowska-Spychalska, D., Kotula, N., Mazurek, G. and Sułkowski, Ł. (2024), "GENERATIVE AI AS SOURCE OF CHANGE OF KNOWLEDGE MANAGEMENT PARADIGM", *Human Technology*, Vol. 20 No. 1, doi: 10.14254/1795-6889.2024.20-1.7.

Levallet, N. and Chan, Y.E. (2019), "Organizational knowledge retention and knowledge loss", *Journal of Knowledge Management*, Vol. 23 No. 1, doi: 10.1108/JKM-08-2017-0358.

Lichtenthaler, U. (2019), "An intelligence-based view of firm performance: Profiting from artificial intelligence", *Journal of Innovation Management*, Vol. 7 No. 1, doi: 10.24840/2183-0606_007.001_0002.

Massingham, P. (2008), "Measuring the impact of knowledge loss: More than ripples on a pond?", *Management Learning*, Vol. 39 No. 5, doi: 10.1177/1350507608096040.

Massingham, P.R. (2018), "Measuring the impact of knowledge loss: a longitudinal study", *Journal of Knowledge Management*, Emerald Publishing, Vol. 22 No. 4, pp. 721–758, doi: 10.1108/JKM-08-2016-0338.

Nonaka, I. and von Krogh, G. (2009), "Tacit knowledge and knowledge conversion: Controversy and advancement in organizational knowledge creation theory", *Organization Science*, Vol. 20 No. 3, doi: 10.1287/orsc.1080.0412.

Nonaka, I. and Takeuchi, H. (1995), "Knowledge-Creating Company", *Knowledge-Creating Company*, No. December 1991.

Polanyi, M. (1966), "The tacit dimension, knowledge in organizations", *L. Prusak*, *Ed*.

Rezaei, M. (2025), "Artificial intelligence in knowledge management: Identifying and addressing the key implementation challenges", *Technological Forecasting and Social Change*, Vol. 217, doi: 10.1016/j.techfore.2025.124183.

Westphaln, K.K., Regoeczi, W., Masotya, M., Vazquez-Westphaln, B., Lounsbury, K., McDavid, L., Lee, H.N., *et al.* (2021), "From Arksey and O'Malley and Beyond: Customizations to enhance a team-based, mixed approach to scoping review methodology", *MethodsX*, Elsevier, Vol. 8, p. 101375, doi: 10.1016/J.MEX.2021.101375.

Zack, M.H. (1999), "Developing a knowledge strategy", *California Management Review*, No. 3, doi: 10.2307/41166000.

Collective and Artificial Intelligence in the Challenge of Happy Degrowth

Marianna Mancino, Maria Zifaro, Giuseppe Ambrosio, Giovanni Spatola, Stefania D'Aprile
Mercatorum University, Rome, Italy

Abstract

The current context is marked by systemic, ecological, social, and economic crises. These are accompanied by an increasing acceleration of technological change. In this scenario, the urgent need emerges to rethink development models based on unlimited economic growth. The degrowth paradigm proposes an alternative orientation centered on principles of sufficiency, equity, and collective well-being. It challenges dominant productivist logics and promotes new forms of social organization. At the same time, artificial intelligence is emerging as a powerful infrastructure capable of influencing organizational and decision-making processes in contemporary societies, reshaping the relationship between technology, work, and society in significant and long-term ways. This contribution explores the joint role of collective intelligence (CI) and artificial intelligence (AI) in supporting scenarios of convivial degrowth. It adopts a theoretical-exploratory and interdisciplinary approach based on the analysis of existing literature and documented case studies. The findings show that AI, if designed and governed according to ethical, inclusive, and democratic principles, can facilitate distributed coordination of activities, knowledge sharing, and collective participation in decision-making processes, thus strengthening local resilience and cooperation. Within this perspective, CI emerges as the cognitive engine of community cooperation and co-creation. Time is understood as a fundamental resource that must be redistributed to enable non-market social practices such as care, volunteering, and participatory deliberation. The paper proposes the Cooperative Socio-Technical Regeneration Model (CSTRM), a theoretical framework that integrates CI, AI, time, and sufficiency as key elements of a socio-ecological transformation oriented toward collective well-being and long-term sustainability. The model highlights that the relationship between AI and degrowth is not automatic. Rather, it is deeply mediated by design, institutional, and cultural choices. The contribution aims to enrich the debate on socio-ecological transition pathways, while acknowledging the limitations related to the theoretical nature of the analysis and the need for future empirical validation, as well as further comparative research across different territorial contexts.

Keywords: Collective Intelligence; Artificial Intelligence; Social Innovation; Happy Degrowth; Collective well-being.

Paper type: Academic Research Paper

1 Introduction

Contemporary transformations, marked by ecological, social, and economic crises, challenge the centrality of growth as the organizing principle of modern societies (Hickel, 2020; Kallis et al., 2018). The degrowth paradigm proposes a reconfiguration of socio-economic systems oriented toward sufficiency, quality of life, and the regeneration of socio-ecological systems (Latouche, 2014; Raworth, 2018; Sachs, 2024). At the same time, artificial intelligence is emerging as a socio-technical infrastructure capable of shaping decision-making and organizational processes, whose impact depends on development contexts and governance logics (Elish and Boyd, 2018; Floridi et al., 2018).

Within this framework, the paper addresses the following research question: how can the interaction between collective intelligence (CI) and artificial intelligence (AI) enable degrowth-oriented pathways, and what are the implications for community structures?

The aim is to develop an integrated understanding of CI and AI as cognitive and socio-technical infrastructures serving an economy of sufficiency. The article adopts a theoretical-exploratory and interdisciplinary approach, based on a narrative literature review and qualitative case study analysis, in order to interpret the relationships between technology and social cooperation.

The main outcome is the Cooperative Socio-Technical Regeneration Model (CSTRM), which conceptualizes CI as the cognitive driver of cooperation, AI as a supporting infrastructure, time as a political resource, and sufficiency as a normative principle. The originality lies in integrating CI, AI, and degrowth within a unified framework, offering an alternative perspective on technological innovation as a lever for collective well-being. Limitations include the theoretical nature of the analysis and the limited generalizability of the cases. Managerial implications concern the design of digital systems oriented toward cooperation, transparency, and participation.

2 Theoretical Background

2.1 Collective Intelligence

Collective intelligence (CI) can be broadly defined as the emergent capacity of a group to produce solutions, decisions, or knowledge that are qualitatively superior to those generated by individuals alone (Wechsler, 1971). It arises from dynamic processes of interaction, communication, and knowledge sharing, rather than from individual attributes.

Lévy conceptualizes CI as "a form of universally distributed intelligence, constantly enhanced, coordinated in real time, and resulting in the effective mobilization of skills" (Lévy, 1997/1994, pp. 13–15). From this perspective, intelligence becomes an expression of the ability of social and organizational contexts to recognize and activate it.

Converging evidence from the "wisdom of crowds" literature shows that heterogeneous groups can outperform individuals when diversity, independence, and decentralization are preserved (Surowiecki, 2005). Similarly, Woolley et al. (2010) identify a collective intelligence factor (c), linked to social sensitivity and balanced participation within groups.

Nevertheless, aggregation alone does not guarantee the emergence of CI. As highlighted by Le Bon (1900), unstructured crowds may produce irrational and unreflective outcomes. CI instead develops within contexts characterized by coordination, structured interaction, and trust. In this regard, cooperation and cognitive diversity are key determinants. High-performing groups are those able to dynamically regulate diversity in relation to task demands, expanding it during exploratory phases and narrowing it during coordination phases, while maintaining internal coherence (Lix et al., 2022; Karadzhov et al., 2024). This dynamic explains the strong association between CI and processes of co-creation and social innovation, where active community involvement plays a crucial role in addressing complex challenges (Mulgan, 2006; Manzini, 2015).

Furthermore, the effectiveness of CI depends on enabling conditions. Technological and organizational infrastructures support coordination, information flows, and the integration of contributions. The "collective intelligence genome" framework, in particular, emphasizes how participation structures, motivational drivers, outputs, and organizational mechanisms shape collaborative dynamics (Malone et al., 2009). Overall, CI can be regarded as a strategic resource for learning, adaptation, and cooperation, contingent upon contexts that enable trust, knowledge sharing, and effective coordination (Broadbent and Gallotti, 2015).

2.2 AI as socio-technical infrastructure

Artificial intelligence (AI) has traditionally been defined as the capacity of machines to perform tasks requiring human-like intelligence, including learning, reasoning, and decision-making (Boden, 2016; Russell and Norvig, 2021). More recently, however, AI has been reconceptualized as a socio-technical infrastructure, whose logics, applications, and effects are deeply embedded within the social contexts of its development and use (Elish and Boyd, 2018).

Within this framework, AI cannot be considered neutral. Its outcomes are shaped by the values embedded in design processes, governance structures, and strategic objectives. In dominant economic paradigms, AI is frequently deployed to enhance efficiency and sustain economic growth, thereby reinforcing extractive dynamics and accelerating material and energy throughput (Hickel, 2020; Hogan, 2024; Jackson, 2014).

Conversely, a growing body of literature emphasizes AI's potential as an enabling technology for social innovation, supporting knowledge sharing, resource coordination, and participatory governance (Floridi et al., 2018; Kim et al., 2022; Manzini, 2015). In this sense, AI may function as an "informational glue" capable of integrating heterogeneous data, identifying patterns, and informing decision-making processes, provided that it is designed according to principles of transparency, inclusiveness, and democratic control (Illich, 1973; Milotay and Sgueo, 2020).

A critical issue concerns the relationship between AI and time. Automation holds the potential to liberate time from repetitive tasks, enabling social participation and care practices. However, without appropriate policy frameworks these gains may instead result in work intensification and the acceleration of everyday life (European Commission, 2024; Morandini et al., 2023).

Against this backdrop, the compatibility between AI and degrowth is contingent upon the criteria used to design and assess technological systems. The sufficiency perspective calls for moving beyond productivity- and efficiency-driven metrics, focusing instead on the forms of life, relationships, and social organization that AI enables (Meyers, 2024; Vetter, 2018; Sachs, 2024).

2.3 Degrowth

The degrowth paradigm constitutes a critical response to the dominance of economic growth as the organizing principle of contemporary societies (Hickel, 2020; Kallis et al., 2018). Rather than prioritizing GDP expansion, it advocates a reconfiguration of socio-economic systems based on sufficiency, social justice, and ecological limits (Latouche, 2014; Sachs, 2024).

Degrowth involves the intentional downscaling of material and energy throughput within planetary boundaries, with the objective of fostering human well-being, equity, and ecological regeneration (Hickel, 2020). Importantly, it should not be equated with recession, but understood as a deliberate and equitable transformation of economic structures (Victor, 2008). This approach calls for decoupling well-being from economic growth and redefining prosperity in qualitative terms, as the capacity to flourish within ecological constraints (Jackson, 2014).

At the core of this paradigm lies the principle of sufficiency, which promotes limiting production and consumption to sustainable levels while distinguishing between essential needs and excessive consumption (Latouche, 2014; Sachs, 2024). Consequently, innovation is assessed not only in terms of efficiency, but also in relation to its contribution to quality of life and collective well-being.

Degrowth is closely intertwined with social innovation, understood as the development of practices and organizational models that address societal needs through new forms of cooperation and collective action (Manzini, 2015; Murray et al., 2010). Within this framework, communities assume an active role in defining needs and co-producing context-sensitive solutions.

Complementary perspectives, such as Doughnut Economics, stress the importance of situating economic activity within a safe and just space, bounded by a social foundation and an ecological ceiling (Raworth, 2018). In the Italian context, the notion of "degrowth for well-being" foregrounds qualitative values such as time, sobriety, and solidarity, advocating a conscious reduction in consumption and a rebalancing of human–nature relations (Pallante, 2011).

Ultimately, the degrowth paradigm calls for a critical reassessment of technology. Innovation is not inherently beneficial, but must be evaluated according to its coherence with sufficiency, equity, and sustainability, avoiding the reproduction of extractive logics and structural inequalities (Berendt, 2019; Floridi et al., 2018; Hogan, 2024; Vetter, 2018).

2.4 Time and Leisure in the Challenge of Happy Degrowth

Within post-growth debates, time is increasingly conceptualized not as a residual outcome of economic activity, but as a social and political resource that conditions democratic participation, care practices, and socio-ecological regeneration. In this sense, time acquires a relational dimension and becomes subject to institutional and cultural transformation.

It is the distribution of time that constitutes a key dimension of contemporary inequality. The notion of time poverty captures situations in which, after accounting for paid work, unpaid work, and personal needs, individuals are left with insufficient discretionary time. This condition has significant implications for both well-being and social participation (Kalenkoski et al., 2011; OECD, 2020), as only a limited portion of time can be classified as leisure.

To clarify this distinction (Figure 1), it is useful to differentiate between residual time and leisure in a strict sense. Residual time includes all time remaining after paid work. It also includes necessary activities, such as unpaid work (care of the house and family) and personal care (sleep, nutrition, and other essential functions).

While essential, these activities are not fully discretionary, since they are constrained by physiological needs and domestic responsibilities. Authentic leisure, therefore, represents only a subset of time off, consisting of discretionary time available for social interaction, creativity, and personal fulfilment (Haworth and Lewis, 2005; OECD, 2020).

Time also has a significant ecological dimension. Accelerated temporal regimes tend to promote time-saving behaviours, often replacing slower, low-carbon practices, such as self-production, with more energy and resource-intensive alternatives (Smetschka et al., 2023).

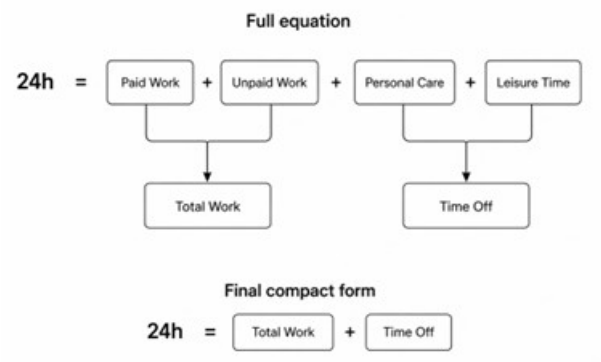


Figure 1. Daily time allocation
Source: Authors' own elaboration based on OECD (2020) data, with AI graphical support

In line with the social acceleration theory, increases in technical and productive speed may coexist with growing perceptions of time scarcity (Rosa, 2003). Although efficiency gains should theoretically liberate time, within profit-driven systems they are frequently absorbed through intensified workloads and expanding demands (Alcott, 2005; Rosa, 2003).

In degrowth scenarios, working time reduction plays a central role, both as a mechanism for redistributing employment and as a means of freeing time for socially valuable activities (Kallis et al., 2018; Lange, 2018).

Empirical evidence indicates that reduced working hours are associated with lower ecological footprints and carbon emissions (Knight et al., 2013), as well as improvements in well-being, including reduced stress and better work–life balance (Kelly et al., 2022). From this perspective, time emerges as a structural condition enabling or constraining participation and cooperation. Happy degrowth can thus be understood as a project of temporal reconfiguration, aimed not only at increasing available time, but at reclaiming it from accumulation logics and redirecting it toward shared well-being.

3 Methodology

The present study adopts an interdisciplinary, theoretical-exploratory approach to explore how the interaction between collective intelligence (CI) and artificial intelligence (AI) may contribute to degrowth pathways. The objective is not to establish causal relationships but to develop an interpretative framework grounded in the combined analysis of literature and illustrative cases.

The research is organized into two phases. First, a narrative review of the literature identifies key conceptual dimensions, including CI, AI as a socio-technical infrastructure, degrowth, and the role of time in post-growth societies. Second, a qualitative analysis of case studies is conducted. Cases are selected through purposive sampling to highlight configurations where CI–AI interactions are consistent with degrowth principles. These cases are not intended to be representative; rather, they are interpreted through the lens of the theoretical framework.

The analysis reveals recurring patterns that inform the development of the Cooperative Socio-Technical Regeneration Model (CSTRM), conceived as an interpretative framework outlining the conditions under which CI and AI can support degrowth. This study presents several limitations. The narrative review is selective and non-exhaustive. The qualitative case analysis is illustrative and does not allow generalization. Moreover, some contexts are evolving. Despite these limitations, the methodology offers an exploratory contribution that advances theoretical insight and supports future research.

4 Results and Discussion

4.1 CI and AI: Illustrative cases

This section examines a selection of cases in which interactions between CI and AI are associated with practices aligned with degrowth. The cases are not intended to be exhaustive. Rather, they serve to illustrate how CI and AI function as interdependent components of socio-technical systems in which human and computational capabilities co-evolve (Elish and Boyd, 2018; Malone et al., 2009).

Within these systems, CI arises from the capacity of heterogeneous actors to mobilize distributed knowledge, which is aggregated and translated into coordinated action (Benkler, 2013; Broadbent and Gallotti, 2015;

Malone et al., 2009). Digital platforms play a key role in enabling this process. AI operates as a complementary layer, supporting the organization of information and facilitating alignment between needs and available resources.

A first set of cases concerns distributed coordination under conditions of fragmentation and informational asymmetry. The Kudu agricultural marketplace in Uganda exemplifies this dynamic. Through an SMS-based platform, farmers and buyers exchange information on prices, quantities, and locations, while AI-based matching algorithms organize these data, incorporating transport costs (Newman et al., 2018; Shi et al., 2020).

However, the case reveals structural limitations. Data reliability issues and logistical constraints necessitate hybrid coordination mechanisms and ongoing human intervention. Furthermore, reliance on external inputs undermines long-term sustainability (Shi et al., 2020).

Despite these constraints, the case demonstrates that the primary value lies not in automation but in enabling localized cooperation. AI enhances rather than replaces human agency, facilitating the alignment of supply and demand as a form of socio-technical innovation (Manzini, 2015).

A comparable logic is evident in the Olio platform, which connects citizens, volunteers, and organizations for the redistribution of surplus food and goods. CI emerges through multi-level participation, where users share information, identify local availability, and respond to community needs. Volunteers, known as Food Waste Heroes, act as coordinating nodes within the network (Cahn, 2000), organizing collection and redistribution processes.

AI supports these interactions through image recognition, automated content generation, moderation, and matching functions (Olio, 2023). The platform also adopts low-energy computational models and promotes circular economy practices through partnerships (Olio, n.d.). Nevertheless, tensions arise. The system depends heavily on unpaid labor while generating potential economic value, raising concerns about the distribution of value and possible misalignment with degrowth principles (Benkler, 2013).

Beyond material coordination, other cases highlight the social recognition of care and everyday support. In these contexts, CI is expressed through trust, reciprocity, and cooperative practices embedded in communities (Manzini, 2015). Time banks represent a paradigmatic example. Based on egalitarian exchange, they assign equal value to all activities, with one hour of work corresponding to one credit (Cahn, 2000).

These systems foreground time as a central resource and make visible forms of labor typically excluded from market valuation (OECD, 2020). However, digital implementations may introduce distortions. The TimeRepublik platform illustrates a shift toward short-term, visibility-driven interactions aligned with self-promotion and market logics (Arcidiacono, 2016), thereby weakening collective orientations.

Mutual aid networks developed during the COVID-19 pandemic provide a more consistent alignment with degrowth principles. Local groups mobilized resources to address concrete needs, including food distribution and assistance to vulnerable populations (Chevée, 2022; Travlou, 2020). Digital tools facilitated coordination without displacing social relations, while trust and proximity remained central.

These cases demonstrate that lightweight digital infrastructures can support co-production as a core organizing principle (Cahn, 2000). Nonetheless, challenges persist, including cultural barriers, trust deficits, safety concerns, and stigma associated with requesting support (Verma et al., 2024).

Citizen science initiatives represent another domain of CI–AI interaction. In coral reef monitoring projects in the Pacific, local communities collaborate with digital tools to assess ecosystem health (Cui and Yasseri, 2024; Irwin, 1995). In resource-constrained environments, these practices assume a community-based configuration (Beery, 2025; Berkes, 2004).

Participants collect data using accessible technologies such as smartphones and underwater cameras. CI manifests in collective observation and interpretation grounded in local ecological knowledge. AI enhances these processes through automated image analysis and species classification, enabling shared knowledge production and environmental decision-making (Beery, 2025; Berkes, 2004). However, infrastructural limitations, funding dependencies, and reliance on volunteer labor may constrain sustainability.

Finally, CI–AI interactions are evident in deliberative democracy processes. Digital platforms enable various forms of e-participation, from information access to active engagement in decision-making (Macintosh, 2004). CI is expressed through the formation of shared judgments via public deliberation (Cohen, 2007). AI contributes by supporting information synthesis, classification, and visualization, helping identify patterns of consensus and disagreement (Cui and Yasseri, 2024; Fontes et al., 2024). The Decidim platform illustrates this dynamic. As an open-source infrastructure, it enables citizens to propose, debate, and vote on public policies (Aragón et al., 2017; García et al., 2023).

The resulting value lies in enhancing decision quality and legitimacy through participation (Berendt, 2019; Raskin et al., 2002). However, these outcomes depend on governance. Without transparency and inclusiveness, participation risks remaining symbolic, limiting democratic impact (Berendt, 2019).

Overall, these cases show that CI–AI interactions are shaped by socio-technical configurations and governance structures, enabling either cooperative and participatory outcomes or trajectories inconsistent with degrowth.

4.2 Toward a cooperative Socio-Technical Regeneration Model (CSTRM)

The cases discussed demonstrate that the interaction between CI and AI is contingent upon its configuration within specific social and institutional settings. Technology is not neutral but embedded in socio-technical systems that shape both its functions and outcomes (Elish and Boyd, 2018; Berendt, 2019). As such, the alignment between AI and degrowth is not inherent, but must be constructed through deliberate design and governance choices.

In response, this paper introduces the Cooperative Socio-Technical Regeneration Model (CSTRM) as a cyclical, non-linear conceptual framework (Figure 2). While the model does not follow a causal sequence, a conceptual entry point can be identified to facilitate its interpretation.

Figure 2. Interpretative Framework of the CSTRM
Source: Authors' own elaboration with AI-assisted graphical support

The model is grounded in the principles of sufficiency and degrowth, which provide its normative orientation. These principles shift the focus from production expansion to improving quality of life and reducing throughput, defined as the flow of energy and materials through the economic system, from resource extraction to production and waste (Kallis et al., 2018).

This orientation leads to a redistribution of social time, including through reduced working hours, with positive effects on well-being and work–life balance (Haraldsson and Kellam, 2021; Schor et al., 2022). Time freed from paid labor becomes available time, partially removed from accumulation logics, and constitutes a key condition for degrowth strategies (Hickel, 2020). The availability of time enables participation, care, and cooperation. In this context, CI acts as a cognitive engine for addressing complex problems, supporting deliberation and coordination (Cui and Yasseri, 2024; Malone et al., 2009), and is sustained by co-creation practices and participatory platforms.

A central role is played by commons, defined as shared resources collectively governed according to community-defined rules (Benkler, 2013; de Rosnay, 2020). Commons provide the foundation for integrated CI–AI configurations. Open infrastructures, data, and knowledge foster cooperation and experimentation, offering alternatives to proprietary models (Benkler, 2013).

AI functions as a socio-technical support when designed to incorporate the social contexts and decision-making processes in which it operates (Elish and Boyd, 2018). In this way, it supports inclusive deliberation, traceability, coordination, and co-production, contributing to the construction of a political "we" (García et al., 2023). Its impact depends on governance arrangements and alignment with equity and sufficiency goals.

These dynamics give rise to regenerative communities: adaptive socio-ecological systems characterized by cooperation, informal social relations, and the capacity to co-evolve with their environments (Gibbons, 2020).

Such communities represent the progressive outcome of the interaction between normative orientation, time availability, CI, and socio-technical support.

The CSTRM thus highlights that CI–AI collaboration can generate degrowth-oriented social innovation only when guided by sufficiency principles. Time emerges as a transversal common good enabling participation and transformation.

4.3 Theoretical and managerial implications

This paper advances the socio-ecological transition debate by conceptualizing CI and AI as interdependent elements of socio-technical infrastructures capable of supporting degrowth-oriented collective processes. Within this perspective, time emerges as a critical political resource that enables participation, care, and cooperative action (Haworth and Lewis, 2005; OECD, 2020; Rosa, 2003).

The proposed Cooperative Socio-Technical Regeneration Model (CSTRM) provides an interpretative framework connecting normative principles, time availability, CI activation, the situated deployment of AI, and the formation of regenerative communities (Gibbons, 2020).

Findings indicate that the relationship between CI, AI, and degrowth is contingent rather than deterministic. It is shaped by socio-institutional contexts and design choices, reinforcing the argument that AI is not neutral (Berendt, 2019; Vetter, 2018). Its potential contribution depends on governance structures and design principles that prioritize transparency and openness (Benkler, 2013; de Rosnay, 2020). While AI can enable cooperation, sharing, and regenerative practices (Manzini, 2015; Murray et al., 2010), it may also reproduce extractive dynamics and systemic inequalities (Floridi et al., 2018; Hickel, 2020; Hogan, 2024; Jackson, 2014).

Accordingly, AI should be assessed based on its capacity to support the common good, conceived as a dynamic, context-sensitive process of collective construction involving plural perspectives and inclusive decision-making (Berendt, 2019). The model emphasizes sufficiency as a normative criterion guiding technological development toward collective well-being.

Managerially, the results suggest the need to redefine design and evaluation criteria for digital systems. Participatory and inclusive design approaches should be prioritized to integrate stakeholder perspectives and distributed knowledge (Manzini, 2015; Benkler, 2013; Malone et al., 2009).

AI systems should be framed as cooperative infrastructures, emphasizing transparency, explainability, and collective oversight, while mitigating risks associated with algorithmic opacity (Elish and Boyd, 2018; Illich, 1973; Milotay and Sgueo, 2020; Floridi et al., 2018). Furthermore, developing low-impact technological solutions that are energy-efficient and locally adaptable is essential (Hickel, 2020; Kallis et al., 2018; Lange, 2018).

Public institutions are crucial in enabling sufficiency-oriented ecosystems by supporting open-source initiatives, digital commons, and equitable governance arrangements (Benkler, 2013; de Rosnay, 2020). Finally, integrating time into policy agendas—particularly through working time reduction and redistribution—can enhance participation, care, and cooperation (Kallis et al., 2018; Rosa, 2003).

4.4 Limitations and future research

This study presents several limitations. The theoretical–exploratory approach does not allow for the establishment of causal relationships or the empirical generalization of findings, as it is oriented toward conceptual development and pattern identification. The analyzed cases are qualitative and illustrative and are influenced by contextual factors, limiting the transferability of results to other settings.

Moreover, the narrative literature review may reflect interpretative selectivity, potentially introducing bias. More broadly, qualitative research entails limitations related to subjectivity and the difficulty of testing cause–effect relationships. An additional concern relates to the role of AI within socio-technical systems, which is subject to issues of transparency, bias, and context dependency.

Future research could adopt empirical approaches to further refine the CSTRM model. Longitudinal studies may examine the evolution of CI–AI configurations over time, while further investigations should explore governance mechanisms, value distribution, digital commons, and the relationship between time, technology, and social organization.

5 Conclusion

This study addressed the following research question: how can the interaction between collective intelligence (CI) and artificial intelligence (AI) contribute to enabling degrowth pathways, and what implications does this have for the community fabric? To clarify the findings, the question was articulated into two sub-

questions: (a) how can CI–AI interaction support degrowth pathways? and (b) what are the implications for community structures?

The analysis shows that the combination of CI and AI can support degrowth pathways when embedded within socio-technical systems oriented toward sufficiency and cooperation. In this perspective, CI functions as the cognitive engine of collective processes, grounded in participation, diversity, and co-creation. AI operates as an infrastructure that supports coordination, knowledge sharing, and the alignment between needs and resources. However, this contribution depends on design and governance choices that ensure transparency, inclusion, and democratic control.

Adopting sufficiency and degrowth as normative principles fosters a transformation of the community fabric. Under these conditions, communities strengthen their capacity for self-organization, co-production, and shared resource management. The role of commons and time emerges as central, as social and political resources that enable cooperation and care practices. Time, in particular, acts as a key enabling condition, making participation and collective action possible.

Such transformation requires a broader paradigm shift supported by institutional and cultural factors. On the one hand, coherent policies and governance arrangements aligned with sufficiency, equity, and participation are needed. On the other hand, these principles must be progressively internalized through processes of social institutionalization, becoming embedded in shared values and practices. In the absence of such orientation, CI–AI interactions risk reinforcing profit-driven dynamics and reproducing inequalities and exclusion.

The main contribution of this paper lies in the proposal of the Cooperative Socio-Technical Regeneration Model (CSTRM), which interprets the CI–AI relationship as a co-evolutionary process guided by degrowth principles. The model highlights the role of time as a social and political resource and of commons as the foundation for alternative configurations beyond market logics.

Ultimately, the issue assumes a political and cultural dimension, linked to the rethinking of development models and socio-technical infrastructures. In this sense, the CSTRM provides an interpretative lens to orient technological innovation within broader socio-ecological transitions.

6 Ethical Declaration

This research did not involve human or animal subjects, nor did it collect any sensitive data. Therefore, approval from an ethics committee was not required.

7 AI Declaration

A limited use of artificial intelligence tools was made exclusively for language support in translation and for the generation of the figures included in the paper. All ideas, structure, and content are original to the authors, who assume full responsibility for them.

References

Alcott, B. (2005) Jevons' paradox, Ecological economics, 54(1), 9-21.

Aragón, P., Kaltenbrunner, A., Calleja-López, A., Pereira, A., Monterde, A., Barandiaran, X. E., and Gómez, V. (2017) Deliberative platform design: The case study of the online discussions in Decidim Barcelona, In International conference on social informatics (pp. 277-287). Cham: Springer International Publishing.

Arcidiacono, D. (2016) Il tempo nell'economia fondamentale tra sharing economy e personal branding: il caso del Time Banking on line, Sociologia del lavoro: 142, 2, 2016, 98-114.

Beery, S. (2025) People-Centric AI for Conserving Biodiversity, J. Artif. Intell. Aquat. Res.

Benkler, Y. (2013) Commons and Growth: The Essential Role of Open Commons in Market Economies (reviewing Infrastructure: The Social Value of Shared Resources by Brett M. Frischmann), University of Chicago Law Review, 80(3), 12.

Berendt, B. (2019) AI for the Common Good?! Pitfalls, challenges, and ethics pen-testing, Paladyn, Journal of Behavioral Robotics, 10(1), 44-65.

Berkes, F. (2004) Rethinking community-based conservation, Conservation biology, 18(3), 621-630.

Boden, M. A. (2016) AI: Its nature and future, Oxford University Press.

Broadbent, S., and Gallotti, M. (2015) Collective intelligence: How does it emerge. NESTA: London, UK.

Cahn, E. S. (2000) No more throw-away people: The co-production imperative, Edgar Cahn.

Chevée, A. (2022) Mutual Aid in north London during the Covid-19 pandemic, Social Movement Studies, 21(4), 413-419.

Cohen, J. (2007) Deliberative democracy. In Deliberation, participation and democracy: Can the people govern? (pp. 219-236) London: Palgrave Macmillan UK.

Cui, H., and Yasseri, T. (2024) AI-enhanced collective intelligence, Patterns, 5(11).

de Rosnay, D. (2020) Digital commons, Internet Policy Review, 9(4).

Elish, M. C., and Boyd, D. (2018) Situating methods in the magic of Big Data and AI, Communication Monographs, 85(1), 57-80.
European Commission (2024) Special Eurobarometer 554 – Artificial Intelligence and the Future of Work, Directorate-General for Employment, Social Affairs and Inclusion.
Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., ... and Vayena, E. (2018) AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations, Minds and machines, 28(4), 689-707.
Fontes, C., Al Haddad, C., Antoniou, C., and Lütge, C. (2024) Assessing the collective challenges of AI: An applied framework focused on people centeredness and public participation, Urban Governance, 4(4), 245-254.
García, D. L., Calleja-López, A., and Linares-Lanzman, J. (2023) Collective intelligence and digital participatory platforms: Learnings from Barcelona's Decidim, In The Routledge handbook of collective intelligence for democracy and governance (pp. 326-338). Routledge.
Gibbons, L. V. (2020) Moving beyond sustainability: A regenerative community development framework for co-creating thriving living systems and its application, Journal of Sustainable Development, 13(2), 20-52.
Haraldsson, G.D. and Kellam, J., (2021) Going public: Iceland's journey to a shorter working week, Alda and Autonomy.
Haworth, J., and Lewis, S. (2005) Work, leisure and well-being. British Journal of Guidance and Counselling, 33(1), 67-79.
Hickel, J. (2020) Less is more: How degrowth will save the world, Random House.
Hogan, M., and LePage-Richer, T. (2024) Extractive AI, Centre for Media, Technology and Democracy.
Illich, I. (1973) Tools for conviviality, World perspectives 47. New York: Harper and Row.
Irwin, A. (1995) Citizen Science: A Study of People, Expertise and Sustainable Development.
Jackson, T. (2014) Prosperità senza crescita: economia per il pianeta reale, Edizioni Ambiente.
Kalenkoski, C. M., Hamrick, K. S., and Andrews, M. (2011) Time poverty thresholds and rates for the US population, Social Indicators Research, 104(1), 129-155.
Kallis, G., Kostakis, V., Lange, S., Muraca, B., Paulson, S., and Schmelzer, M. (2018) Research on degrowth, Annual review of environment and resources, 43(1), 291-316.
Karadzhov, G., Vlachos, A., and Stafford, T. (2024) The effect of diversity on group decision-making, arXiv preprint arXiv:2402.01427.
Kelly, O., Schor, J., Fan, W., Bezdenezhnykh, T., Gu, G., and Bridson Hubbard, N. (2022), The Four Day Week: Assessing global trials of reduced work time with no reduction in pay: Evidence from Ireland, University College Dublin.
Kim, E., Jang, G. Y., and Kim, S. H. (2022) How to apply artificial intelligence for social innovations, Applied Artificial Intelligence, 36(1), 2031819.
Knight, K., Rosa, E. A., and Schor, J. B. (2013) 12. Reducing growth to achieve environmental sustainability: the role of work hours, Capitalism on trial: Explorations in the tradition of Thomas E. Weisskopf, 187.
Lange, S. (2018) Macroeconomics without growth, Marburg: Metropolis Verlag.
Latouche, S. (2014) La scommessa della decrescita, Feltrinelli editore.
Le Bon, G. (1900) Psychologie des foules, F. Alcan.
Lévy, P. (1997) Collective intelligence: Mankind's emerging world in cyberspace (R. Bononno, Trans.), Perseus Books. (Original work published 1994).
Lix, K., Goldberg, A., Srivastava, S. B., and Valentine, M. A. (2022) Aligning differences: Discursive diversity and team performance, Management Science, 68(11), 8430-8448.
Macintosh, A. (2004) Characterizing e-participation in policy-making, In 37th Annual Hawaii International Conference on System Sciences, 2004. Proceedings of the (pp. 10-pp). IEEE.
Malone, T. W., Laubacher, R., and Dellarocas, C. (2009) Harnessing crowds: Mapping the genome of collective intelligence, MIT Sloan Research Paper.
Manzini, E. (2015) Design, when everybody designs: An introduction to design for social innovation, MIT press.
Meyers, M. (2024) Artificial Intelligence in a degrowth context: A conviviality perspective on machine learning, GAIA-Ecological Perspectives for Science and Society, 33(1), 186-192.
Milotay, N., and Sgueo, G. (2020) Collective intelligence at EU level-Social and democratic dimensions, European Parliamentary Research Service.
Morandini, S. et al., (2023) The impact of artificial intelligence on workers' skills: Upskilling and reskilling in organisations, Informing Science: The International Journal of an Emerging Transdiscipline, 26, pp.39–68.
Mulgan, G. (2006) The process of social innovation, Innovations, 1(2), 145-162.
Murray, R., Caulier-Grice, J., and Mulgan, G. (2010) The open book of social innovation (Vol. 24), London: Nesta.
Newman, N., Bergquist, L. F., Immorlica, N., Leyton-Brown, K., Lucier, B., McIntosh, C., ... and Ssekibuule, R. (2018) Designing and evolving an electronic agricultural marketplace in Uganda, In Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies (pp. 1-11).
OECD (2020) How's Life? 2020: Measuring Well-being, OECD Publishing, Paris.
Olio. (2023) How OLIO works: Share more, waste less [Video], YouTube, Available at https://www.youtube.com/watch?v=KfKVPxTk7y0
Olio. (n.d.) Olio's use of AI, Olio Help Center, Available at https://help.olioapp.com/en/articles/12148938-olio-s-use-of-ai
Pallante, M. (2011) La felicità sostenibile, Rizzoli.
Raskin, P., Banuri, T., Gallopín, G., Gutman, P., Hammond, A., Kates, R., and Swart, R. (2002) Great transition: The promise and lure of the times ahead, Stockholm Environment Institute; Tellus Institute.

Raworth, K. (2018) Doughnut economics: Seven ways to think like a 21st century economist, Chelsea Green Publishing.
Rosa, H. (2003) Social acceleration: ethical and political consequences of a desynchronized high–speed society, Constellations: An International Journal of Critical and Democratic Theory, 10(1).
Russell, S. J., and Norvig, P. (2021) Artificial intelligence: A modern approach (4th ed., Global ed.), Harlow: Pearson.
Sachs, W. (2024) Economia della sufficienza: Appunti per resistere all'Antropocene (V. Dragoni, A. Clò, S. Soglia, and E. Lenzi, trad.), Castelvecchi.
Schor, J. B., Fan, W., Kelly, O., Gu, G., Bezdenezhnykh, T., and Bridson-Hubbard, N. (2022) Assessing global trials of reduced work time with no reduction in pay, Auckland: Four Day Week Global Foundation.
Shi, Z. R., Wang, C., and Fang, F. (2020) Artificial intelligence for social good: A survey, arXiv preprint arXiv:2001.01818.
Smetschka, B., Gaube, V., and Mader, K. (2023) Time to care—Care for time—How spending more time for care than consumption helps to mitigate climate change, Frontiers in Sustainability, 4, 1070253.
Surowiecki, J. (2005) The wisdom of crowds, Knopf Doubleday Publishing Group.
Travlou, P. (2020) Kropotkin-19: A mutual aid response to COVID-19 in Athens, Design and Culture, 13(1), 65-78.
Verma, A., Sridhar, S., Dar, K. F., and Asthana, M. K. (2024) Scope and challenges in the implementation of Time Bank in India: a qualitative study, Discover Global Society, 2(1), 14.
Vetter, A. (2018) The matrix of convivial technology–assessing technologies for degrowth, Journal of cleaner production, 197, 1778-1786.
Victor, P. A. (2008) Managing without growth: Slower by design, not disaster, In Managing Without Growth, Edward Elgar Publishing.
Wechsler, D. (1971) Concept of collective intelligence, American Psychologist, 26(10), 904.
Woolley, A. W., Chabris, C. F., Pentland, A., Hashmi, N., and Malone, T. W. (2010) Evidence for a collective intelligence factor in the performance of human groups, Science, 330(6004), 686-688.

Explainable AI as Decision Support in Automated Assessment: An Exploratory Study

Gabriele Creta, Andrea De Mauro
LUISS Guido Carli University, Rome, Italy

Abstract

This paper investigates how rubric-based Explainable Artificial Intelligence (XAI) outputs influence instructor trust and grading behavior in AI-assisted academic assessment. As AI technologies become increasingly integrated into educational evaluation workflows, understanding the cognitive mechanisms through which instructors interact with AI-generated scores is critical for designing effective human–AI collaboration in grading and for the broader governance of knowledge assets in higher education (Khosravi et al., 2022). Two competing hypotheses motivate the study: explanations may either mitigate automation bias by enabling critical evaluation, or amplify it by providing a rationalization for uncritical acceptance of AI recommendations. We present a within-subject pilot experiment in which university-level evaluators (n = 3) assessed 12 short student responses to Computer Science questions under three conditions: (a) AI-only, where only a numerical AI score was displayed; (b) AI + Explanation, where the AI score was accompanied by a four-criterion rubric breakdown and a textual justification; and (c) Anchoring, where evaluators first graded independently and then were exposed to the AI output with an option to revise. AI grading outputs were generated using a Retrieval-Augmented Generation (RAG) pipeline calibrated against three open-access essay scoring datasets (GRE Analytical Writing, DREsS, and ASAP2) and indexed in a FAISS vector store for similarity retrieval. Preliminary results from this pilot are consistent with rubric-based explanations reducing deviation between human and AI grades (mean |delta| = 1.2 vs. 1.9 for AI-only), doubling the observed revision rate in the anchoring condition (67% vs. 33%), and shifting all observed revisions toward the AI score. Qualitative evidence indicates that evaluators cite specific AI observations when revising with explanations but resist adjustment without them. A discrepancy between self-reported and observed AI influence further suggests a metacognitive blind spot consistent with the automation-bias literature. These preliminary findings are consistent with the hypothesis that structured XAI outputs may enhance evaluator reliance on AI grading systems and reshape evaluative judgment through anchoring mechanisms, with implications for the governance of human–AI assessment workflows in higher education.

Keywords: Explainable Artificial Intelligence; Human–AI Decision-Making; Automated Assessment; Decision Support Systems; Knowledge Governance

Paper type: Academic Research Paper

1 Introduction

Assessment represents a core educational practice through which knowledge is validated, academic standards are enforced, and institutional legitimacy is maintained. Grading decisions do not merely certify learning outcomes; they shape perceptions of fairness, influence trust in educational institutions, and contribute to the governance of academic systems. As artificial intelligence technologies become increasingly integrated into educational environments, assessment has emerged as a particularly critical domain where automated tools promise efficiency gains while simultaneously raising significant ethical, organizational, and governance-related concerns (Kasneci et al., 2023).

The rapid diffusion of Large Language Models (LLMs) within educational environments has introduced new challenges for automated assessment practices, including difficulties in understanding how grading decisions are produced, the emergence of potential biases embedded in model outputs, and a reduced visibility of human judgment in evaluative processes that were traditionally grounded in expert interpretation and contextual awareness (Baker and Hawn, 2021). As a result, educators may develop reduced trust in automated grading systems or, conversely, rely on them inappropriately—either by delegating evaluative responsibility uncritically or by rejecting potentially useful tools due to a lack of interpretability. Both dynamics risk compromising the effective and responsible integration of AI-based assessment systems in educational practice.

Explainable AI (XAI) offers a potential mechanism for promoting calibrated trust rather than blind reliance. In the context of AI-assisted assessment, XAI encompasses techniques that make AI grading decisions interpretable to human evaluators, from feature importance scores to natural language explanations of the reasoning behind an AI grade. Arrieta et al. (2020) provide a comprehensive taxonomy of XAI approaches, emphasizing that the goal of explainability is not complete access to a model's internal reasoning processes, but rather the ability to produce justifications that are understandable and verifiable by domain experts. In educational settings specifically, Khosravi et al. (2022) argue that explainability is essential for maintaining

accountability and trust when AI systems are deployed in contexts where automated decisions have direct consequences for individuals.

However, a parallel concern arises from cognitive psychology. Research on the anchoring-and-adjustment heuristic has established that individuals are susceptible to anchoring effects when making numerical judgments under uncertainty (Tversky and Kahneman, 1974). In the context of AI-assisted grading, an AI-generated score may serve as an anchor that systematically biases the instructor's final grade toward the AI suggestion, regardless of whether the AI assessment is accurate. This phenomenon is related to automation bias, where human decision-makers over-rely on automated systems without critically evaluating their outputs (Parasuraman and Riley, 1997). Explanations may either mitigate this bias by enabling critical evaluation or amplify it by providing a rationalization for uncritical acceptance of AI recommendations (Bansal et al., 2021).

This study addresses these competing hypotheses through a controlled within-subject experiment. Adopting a human-centered perspective, we conceptualize automated grading not as a replacement for educators' judgment, but as a decision-support mechanism integrated into existing assessment practices. In line with a human-centered interpretation of XAI, explainability is operationalized as the ability to justify scoring decisions through explicit, rubric-based criteria that are understandable and verifiable by educators. Specifically, we investigate two research questions:

- ***RQ1:*** *Does the presence of rubric-based XAI explanations increase evaluator reliance on AI grades compared to score-only output?*
- ***RQ2:*** *Do XAI explanations amplify or mitigate anchoring effects when evaluators are exposed to AI grades after forming their own initial assessment?*

2 Theoretical Background

2.1 Automated assessment and AI in education

Automated essay scoring (AES) systems have evolved significantly since their introduction in the 1960s. Contemporary systems leverage deep learning architectures and large language models to evaluate student writing across multiple dimensions including content accuracy, argumentation quality, and language proficiency (Ke and Ng, 2019). Despite improvements in accuracy, adoption in summative assessment contexts remains limited, largely due to instructor skepticism about the ability of AI to capture the nuanced qualitative judgments that characterize expert evaluation (Hockly, 2019). Recent advances in generative AI have introduced new capabilities for rubric-based grading that go beyond holistic scoring: LLMs can generate structured evaluations aligned with predefined criteria, producing not only scores but also detailed justifications that approximate the format of human feedback (Mizumoto and Eguchi, 2023). Yoo et al. (2025) demonstrate the viability of rubric-based essay evaluation using structured datasets such as DREsS, confirming that scoring rubrics can serve as structured knowledge representations that formalize evaluative criteria and make assessment logic explicit.

2.2 Explainable AI, trust, and automation bias

The field of XAI has emerged in response to concerns about the opacity of machine learning models and the need for transparency in high-stakes decision-making contexts. Arrieta et al. (2020) provide a foundational survey of XAI concepts and taxonomies, categorizing approaches by scope, methodology, and explanation type. In educational contexts specifically, Khosravi et al. (2022) identify explainability as a critical enabler for the responsible adoption of AI, arguing that without interpretable outputs, AI-driven educational systems risk undermining both pedagogical objectives and institutional trust.

Trust calibration—the alignment between a user's trust in an AI system and the system's actual reliability—is a central goal of XAI design (Lee and See, 2004). Miscalibrated trust manifests as either under-reliance (rejecting accurate AI outputs) or over-reliance (accepting inaccurate outputs without scrutiny). Explanations have been theorized to promote appropriate reliance by enabling users to evaluate the quality of AI reasoning on a case-by-case basis (Bansal et al., 2021). However, empirical evidence remains mixed: Vasconcelos et al. (2023) find that while explanations can reduce overreliance in some decision-making contexts, they may also increase automation bias by providing a rationalization for uncritical acceptance. Parasuraman and Manzey (2010) provide a comprehensive integration of the automation bias literature, demonstrating that the gap between perceived and actual reliance on automated systems is a robust and persistent phenomenon across professional domains.

2.3 Anchoring effects in evaluative judgment

The anchoring-and-adjustment heuristic, first described by Tversky and Kahneman (1974), refers to the tendency for individuals to insufficiently adjust their estimates from an initial reference value. In assessment contexts, exposure to prior evaluations or suggested scores has been shown to systematically influence subsequent judgments (Mussweiler and Strack, 1999). The introduction of AI-generated grades creates a novel anchoring scenario in which the anchor is produced by an automated system rather than a human peer, raising questions about whether the source of the anchor modulates its influence on evaluator behavior. Few studies have examined anchoring in the specific context of AI-assisted educational assessment, and fewer still have investigated how the presence or absence of AI explanations moderates anchoring strength. This study addresses this gap directly.

3 Research Design and Methodology

3.1 Experimental design

We employed a within-subject experimental design in which each participant evaluated 12 short student responses under three conditions. The within-subject approach controls for individual differences in grading severity, as each participant serves as their own control. The three conditions were: (a) AI-first AI-only (4 tasks), where participants viewed the student response alongside an AI-generated total score with no further explanation; (b) AI-first AI + Explanation (4 tasks), where participants viewed the response alongside the AI score, a criterion-level breakdown, and a textual explanation highlighting key strengths and weaknesses; and (c) Anchoring (4 tasks, 2 with score only and 2 with full explanation), where participants first assigned an initial grade without AI input, then on a subsequent page were shown the AI output and given the option to revise with a brief justification.

3.2 Stimuli: student responses

Twelve short written responses (50–150 words) to fundamental Computer Science questions were collected from undergraduate students in the final year of a Bachelor's degree program. Topics covered core CS concepts including data structures, algorithm complexity, databases, programming paradigms, version control, machine learning, neural networks, and data preprocessing. Responses were reviewed for authenticity; five responses exhibiting characteristics of AI generation (e.g., formulaic analogies, perfectly balanced paragraph structures) were rewritten to match natural student writing style, including realistic imperfections such as minor spelling errors, informal phrasing, and uneven argumentation depth.

3.3 AI grading pipeline

AI grades were generated using a Retrieval-Augmented Generation (RAG) pipeline designed to produce calibrated, rubric-aligned assessment outputs. Three open-access essay scoring datasets were ingested as calibration sources: the GRE Analytical Writing scoring guide with sample essays and reader commentaries at each level (0–6); the DREsS dataset comprising approximately 29,000 student essays scored on three analytic dimensions (Yoo et al., 2025); and the ASAP2 dataset containing approximately 24,000 student essays with holistic scores (1–6). Source texts were segmented into overlapping chunks, embedded using a sentence-level transformer model (all-MiniLM-L6-v2), and indexed in a FAISS vector store for cosine-similarity retrieval. For each student response, the system retrieved the most semantically relevant reference excerpts and injected them into the grading prompt. The Large Language Model used for grading was Claude Opus 4 (Anthropic), operating at low temperature (0.1) to ensure deterministic outputs. All outputs were frozen prior to participant exposure.

3.4 Rubric design

Table 1. Analytic rubric structure

Criterion	Scale	Description
Completeness / Coverage	0–8	Addresses all parts of the question; covers key concepts
Correctness / Conceptual Accuracy	0–8	Factual accuracy; appropriate use of terminology
Reasoning & Structure	0–7	Logical coherence; argument development; organization
Clarity of Language	0–7	Clarity; grammar; readability; conciseness
Total	0–30	Sum of all four criteria

Scoring rubrics were treated as structured knowledge representations that formalize evaluative criteria and make assessment logic explicit. Prior research has shown that rubric-based assessment supports transparency, consistency, and pedagogical effectiveness in grading practices (Yoo et al., 2025). The four-criterion rubric was designed to evaluate both content mastery (Completeness, Correctness) and communicative competence (Reasoning, Language), reflecting the multi-dimensional nature of academic assessment.

3.5 Participants and procedure

Three university-level evaluators participated in the study: two Researchers/Assistant Professors specializing in Artificial Intelligence (10 and 5 years of experience respectively) and one Associate Professor in Data Science (2 years of experience). All held teaching and assessment responsibilities at the university level and were familiar with the Computer Science topics covered. The experiment was administered through a Qualtrics online survey. After providing demographic information, participants completed the 12 evaluation tasks in randomized order. Backward navigation was disabled to prevent modification of initial grades after AI exposure. The survey concluded with a seven-item trust and perception questionnaire (7-point Likert scale). Completion time ranged from 9 to 25 minutes (mean: 17 minutes).

4 Results

4.1 AI-first conditions: effect of explanations on reliance

Table 2. Human grades versus AI scores in AI-first conditions

Condition	Topic	AI	P1	P2	P3	M	Δ
AI-only	Train/Val/Test	25	24	25	22	23.7	−1.3
AI-only	Struct. vs Unstruct.	28	28	27	29	28.0	0.0
AI-only	Relational vs NoSQL	24	25	25	23	24.3	+0.3
AI-only	API	17	25	22	18	21.7	+4.7
AI+Expl	Compiled vs Interp.	26	26	25	24	25.0	−1.0
AI+Expl	Version Control	24	26	24	24	24.7	+0.7
AI+Expl	Overfitting	28	30	27	27	28.0	0.0
AI+Expl	Data Normalization	24	28	24	25	25.7	+1.7

The mean absolute deviation from the AI score was 1.92 points in the AI-only condition and 1.17 points in the AI + Explanation condition, a pattern consistent with rubric-based explanations being associated with greater evaluator reliance on AI grades in this pilot sample. The effect is particularly visible in cases where the AI and human assessments diverge substantially: for the API question (AI score: 17), all three evaluators assigned markedly higher grades in both conditions (18–25), but the deviation was larger in the AI-only condition (mean delta: +4.7) where evaluators lacked access to the AI's reasoning about the response's circularity and superficiality.

4.2 Anchoring analysis: grade revision after AI exposure

Table 3. Anchoring effect: initial grades, AI scores, and revisions

Cond.	Topic	AI	P	Init.	Rev.	Adj.	Dir.
Only	Big O	20	P1	24	23	−1	Toward AI
Only	Big O	20	P2	24	—	0	—
Only	Big O	20	P3	16	18	+2	Toward AI
Only	Neural Net.	22	P1	27	—	0	—
Only	Neural Net.	22	P2	10	—	0	—
Only	Neural Net.	22	P3	22	—	0	—
Expl	API	17	P1	26	24	−2	Toward AI
Expl	API	17	P2	22	20	−2	Toward AI
Expl	API	17	P3	22	21	−1	Toward AI
Expl	Sup./Unsup.	28	P1	30	29	−1	Toward AI
Expl	Sup./Unsup.	28	P2	27	—	0	—
Expl	Sup./Unsup.	28	P3	29	—	0	—

In the AI-only anchoring condition, 2 out of 6 evaluator-task pairs (33%) resulted in grade revision. In the AI + Explanation condition, 4 out of 6 (67%) resulted in revision. All six observed revisions moved in the direction

of the AI score, confirming the presence of an anchoring effect consistent with the predictions of Tversky and Kahneman (1974). The mean adjustment magnitude was 1.5 points for AI + Explanation revisions compared to 0.5 for AI-only, suggesting that explanations not only increase the likelihood of revision but also produce larger adjustments toward the AI anchor.

4.3 Qualitative evidence: the role of explanations in revision reasoning

Analysis of open-ended revision justifications reveals a qualitative difference in reasoning between conditions. In the AI + Explanation condition, evaluators referenced specific observations from the AI explanation to justify their revisions. Participant 1 revised the API question from 26 to 24, stating that the AI "made me think about the shallowness and the lack of key aspects." Participant 2 revised the same question from 22 to 20, noting that "the explanation suggested additional limitations, to which I agree." For the supervised/unsupervised learning question, Participant 1 revised from 30 to 29, acknowledging that "the Netflix example is hybrid, not purely unsupervised."

In the absence of explanations, evaluators expressed explicit resistance to adjustment. Participant 1 stated: "Without explanation I just stick to my grade," providing direct evidence that rubric-based explanations serve as a cognitive mechanism enabling evaluators to identify specific grounds for revision, whereas a numerical score alone is insufficient to overcome confirmation of one's own judgment. This finding is consistent with Bansal et al.'s (2021) observation that explanations enable complementary team performance by providing decision-makers with information they can independently verify.

4.4 Trust and perception questionnaire

Table 4. Trust and perception questionnaire (1–7 Likert scale)

Statement	P1	P2	P3	M
AI grade was accurate and reliable	5	5	5	5.0
Confident in my final grade	6	6	6	6.0
AI influenced my evaluation	5	3	2	3.3
AI made evaluation faster	6	5	5	5.3
Understood AI grading logic	4	7	5	5.3
AI makes evaluation fairer	7	7	4	6.0
Would use AI grading regularly	7	7	3	5.7

The most notable finding is the discrepancy between self-reported and observed AI influence. Evaluators reported that the AI had limited influence on their evaluation (mean: 3.3/7, below the scale midpoint), yet the behavioral data show that 50% of all anchoring task pairs resulted in grade revision, with 100% of revisions moving toward the AI score. This pattern is consistent with the automation bias literature documenting a gap between perceived and actual reliance on automated systems (Parasuraman and Manzey, 2010). At the same time, evaluators expressed high confidence in their final grades (6.0/7) and strong agreement that AI makes evaluation fairer (6.0/7), suggesting that they value AI as a decision-support tool even while underestimating its influence on their own judgment.

5 Discussion

The results of this exploratory study provide preliminary evidence addressing both research questions and contribute to the emerging literature on human–AI collaboration in educational assessment.

Regarding RQ1, the AI + Explanation condition produced closer alignment between human and AI grades than the AI-only condition (mean |delta| = 1.17 vs. 1.92), a pattern consistent with rubric-based XAI outputs being associated with greater evaluator reliance on AI assessments. This finding is consistent with the trust calibration framework proposed by Lee and See (2004), which predicts that transparency in AI decision-making facilitates appropriate reliance. It also aligns with Khosravi et al.'s (2022) argument that explainability is essential for productive human–AI collaboration in educational contexts: structured explanations appear to provide the interpretive scaffolding that enables evaluators to engage meaningfully with AI outputs rather than simply accepting or rejecting them.

Regarding RQ2, the anchoring analysis suggests that, in this pilot sample, explanations may amplify rather than mitigate anchoring effects. Evaluators were twice as likely to revise their initial grade when presented with an explanation compared to a score alone, and all revisions moved toward the AI score. This finding challenges the optimistic view that XAI necessarily promotes critical evaluation. Instead, structured explanations may function as persuasive arguments that facilitate acceptance of the AI perspective, particularly when the

explanation identifies specific weaknesses that the evaluator had not independently considered. This interpretation is consistent with Vasconcelos et al.'s (2023) finding that explanations can increase overreliance in certain decision contexts.

The qualitative data provide a mechanistic account: when explanations are present, evaluators engage in a form of deliberative anchoring, evaluating the plausibility of specific AI observations and adjusting when they find the reasoning convincing. When explanations are absent, evaluators lack a cognitive basis for adjustment and default to their initial judgment. The trust questionnaire reveals an important metacognitive blind spot: evaluators perceive AI as accurate and fair but underestimate its influence, suggesting that AI-assisted grading may introduce systematic biases that evaluators are not consciously aware of. This has significant implications for the governance of AI-enabled assessment systems, aligning with broader concerns about accountability and transparency in algorithmic decision-making (Baker and Hawn, 2021).

5.1 Implications for practice

These findings carry practical implications for the design of AI-assisted assessment systems. First, structured explanations should be considered a design requirement: they meaningfully increase evaluator engagement with AI outputs and enable more informed grading decisions. Second, institutions deploying AI grading tools should implement awareness training to help evaluators recognize anchoring effects and automation bias. Third, assessment workflows should consider the sequencing of human and AI evaluation carefully, as the order of exposure fundamentally shapes the interaction between human judgment and AI suggestion. Finally, rubric-based explanations may serve a dual function as both a transparency mechanism and a calibration tool, helping to align grading standards across multiple evaluators.

6 Future Research Directions

As an exploratory pilot, this study opens several promising avenues for further investigation. The sample size (n = 3) is consistent with the proof-of-concept nature of this work: the within-subject design provides multiple observations per participant and supports the identification of behavioral patterns, while inferential statistical analysis is reserved for a planned full-scale extension with a substantially larger and more diverse evaluator pool. A second extension concerns disciplinary scope: all participants specialized in AI or Data Science, and replication with evaluators from non-technical disciplines (e.g., humanities, social sciences) will help establish whether the observed patterns generalize across academic fields. Third, the present pilot focused on short-answer responses in Computer Science; subsequent work will examine longer-form assessments and a wider range of disciplines, where rubric structure and evaluative complexity differ substantially. Fourth, the AI pipeline employed a single model and configuration, and future iterations will benchmark alternative LLMs and retrieval strategies to assess the robustness of the observed effects. Finally, while this pilot intentionally compared AI-only and AI + Explanation conditions, a follow-up design will incorporate a pure human-only baseline to quantify the marginal contribution of AI assistance to overall assessment quality. Taken together, these directions define a structured research agenda that builds directly on the present findings.

7 Conclusion

This study provides preliminary evidence that rubric-based Explainable AI outputs may meaningfully shape evaluator behavior in AI-assisted academic grading. In this pilot sample, the presence of structured explanations was associated with reduced deviation between human and AI grades, a higher likelihood of grade revision after AI exposure, and revisions that consistently moved toward the AI score. Qualitative evidence further suggests that explanations provide the cognitive scaffolding that enables evaluators to engage critically with AI reasoning, while the absence of explanations leads to resistance to adjustment. These findings contribute to the broader discourse on human–AI collaboration and knowledge governance in educational organizations, supporting the view that explainability is not an intrinsic property of the model alone, but rather an emergent characteristic shaped by the provision of contextual information, evaluative guidance, and structured justification. As AI tools become increasingly prevalent in educational assessment, designing systems that promote calibrated trust rather than uncritical adoption will be essential for preserving the integrity and fairness of academic evaluation.

8 Ethics declaration

This research involved voluntary participation in an anonymous online survey. No personal identifying information was collected beyond professional role and years of experience. Participation was voluntary and participants could withdraw at any time.

9 AI declaration

AI tools (Claude, Anthropic) were used in the following capacities: (1) as the grading engine within the RAG pipeline to generate AI assessment outputs used as experimental stimuli; (2) as an editorial aid for manuscript preparation. All substantive research design, data collection, analysis, and interpretation were conducted by the authors.

References

Arrieta, A.B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., et al. (2020) 'Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI', *Information Fusion*, 58, pp. 82–115.

Baker, R.S. and Hawn, A. (2021) 'Algorithmic bias in education', *International Journal of Artificial Intelligence in Education*, 32, pp. 1052–1092.

Bansal, G., Wu, T., Zhou, J., Fok, R., Nushi, B., Kamar, E., Ribeiro, M.T. and Weld, D.S. (2021) 'Does the whole exceed its parts? The effect of AI explanations on complementary team performance', *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1–16.

Hockly, N. (2019) 'Automated writing evaluation', *ELT Journal*, 73(1), pp. 82–88.

Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., et al. (2023) 'ChatGPT for good? On opportunities and challenges of large language models for education', *Learning and Individual Differences*, 103, 102274.

Ke, Z. and Ng, V. (2019) 'Automated essay scoring: A survey of the state of the art', *Proceedings of the 28th International Joint Conference on Artificial Intelligence*, pp. 6300–6308.

Khosravi, H., Buckingham Shum, S., Chen, G., Conati, C., Tsai, Y.-S., Kay, J., Knight, S., Martinez-Maldonado, R., Sadiq, S. and Gašević, D. (2022) 'Explainable artificial intelligence in education', *Computers & Education: Artificial Intelligence*, 3, 100074.

Lee, J.D. and See, K.A. (2004) 'Trust in automation: Designing for appropriate reliance', *Human Factors*, 46(1), pp. 50–80.

Mizumoto, A. and Eguchi, M. (2023) 'Exploring the potential of using an AI language model for automated essay scoring', *Research Methods in Applied Linguistics*, 2(2), 100050.

Mussweiler, T. and Strack, F. (1999) 'Hypothesis-consistent testing and semantic priming in the anchoring paradigm', *Journal of Experimental Social Psychology*, 35(2), pp. 136–164.

Parasuraman, R. and Manzey, D.H. (2010) 'Complacency and bias in human use of automation: An attentional integration', *Human Factors*, 52(3), pp. 381–410.

Parasuraman, R. and Riley, V. (1997) 'Humans and automation: Use, misuse, disuse, abuse', *Human Factors*, 39(2), pp. 230–253.

Tversky, A. and Kahneman, D. (1974) 'Judgment under uncertainty: Heuristics and biases', *Science*, 185(4157), pp. 1124–1131.

Vasconcelos, H., Joshi, M., Bansal, G., Kamar, E. and Weld, D.S. (2023) 'Explanations can reduce overreliance on AI systems during decision-making', *Proceedings of the ACM on Human-Computer Interaction*, 7(CSCW1), pp. 1–45.

Yoo, H., Han, J., Ahn, S.-Y. and Oh, A. (2025) 'DREsS: Dataset for rubric-based essay scoring on EFL writing', *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (ACL)*, pp. 13439–13454.

Appendix A. Example of AI grading output (AI + Explanation condition)

The following excerpt illustrates the structure of the AI grading output presented to evaluators in the AI + Explanation condition. It comprises (i) the question prompt, (ii) the verbatim student response, (iii) the AI-suggested score with rubric breakdown, and (iv) the textual explanation generated by the RAG pipeline.

Question (Q25). Explain what data normalization means in the context of data preprocessing. Why is it important before applying machine learning algorithms?

Student response. "Data normalization, used for preprocessing, is a way to convert numerical value scale of a dataset into a standard interval typically between 0 and 1. It ensures that all the features contribute equally to the model and that very large nominal value doesn't influence improperly the training phase. It is used to accelerate the training because algorithms can find optimal solution faster."

AI suggested score: 24/30. Completeness: 6/8 | Correctness: 7/8 | Reasoning: 5/7 | Language: 6/7.

Explanation. The answer covers the core definition (scaling to a standard interval) and two important reasons for normalization (equal feature contribution, faster training convergence), addressing the main parts of the question. Content is accurate, though "nominal value" appears to be a misuse — the student likely means "numerical value" or is confusing "nominal" (categorical) with "large-magnitude." The 0-to-1 range is correct

for min-max normalization but the answer doesn't distinguish it from standardization (z-score), which would strengthen completeness. The answer is concise and functional but could develop the reasoning further — for instance, explaining which algorithms are particularly sensitive to feature scaling (e.g., KNN, gradient descent-based methods) versus which are not (e.g., tree-based methods).

How Healthcare Artificial Intelligence Startups Impact Clinical Processes in the Healthcare Ecosystem

Nunzia Carbonara, Davide De Vita, Umberto Panniello, Elisabetta Vurchio
Polytechnic University of Bari, Italy

Abstract

Artificial Intelligence (AI) is increasingly recognized as a transformative force in healthcare, with the potential to enhance clinical efficiency, reduce costs, and enable more personalized care pathways. In parallel, AI has stimulated the emergence of numerous startups developing innovative healthcare solutions. However, existing research has largely focused on their technological features and business models, overlooking their broader systemic role. AI startups are entering healthcare ecosystems as new actors, establishing novel relationships with providers, patients, regulators, and technology suppliers, and potentially reshaping structures, coordination mechanisms, and value creation processes. This study addresses this gap by examining how European healthcare AI startups contribute to the redesign of clinical processes and the reconfiguration of value creation within healthcare ecosystems. Building on the ecosystem-as-structure perspective and a role-based view of ecosystem participation, healthcare is conceptualized as a complex system characterized by interdependencies among heterogeneous actors. Within this framework, AI startups are not merely technology providers but actors capable of altering activity flows, redistributing roles, and generating new connections among stakeholders. Adopting a service and ecosystem perspective, the study emphasizes that value does not reside in technological solutions per se but emerges through their use and integration within specific contexts. AI solutions may generate value for patients—through improved accessibility, personalization, and continuity of care—and for providers—through enhanced decision-making and efficiency. At the same time, they may also produce value destruction when they create workflow frictions, usability issues, or misalignment with clinical routines. Focusing on the European context, characterized by strong regulation and data governance constraints, the study employs a mixed-method approach. First, a large dataset of healthcare AI startups is analyzed using a hybrid natural language processing technique to classify firms into core clinical processes (diagnosis, treatment, monitoring/follow-up) and support processes (drug discovery, administrative support, and healthtech infrastructure). Second, user-generated reviews are examined to assess perceived value creation and value destruction across clinical phases and stakeholder groups. Preliminary findings suggest that AI startups play a dual role within healthcare ecosystems, intervening both directly in clinical activities and indirectly through enabling infrastructures and support functions. Their impact extends beyond isolated applications, depending on how they integrate into existing workflows and reshape interactions among ecosystem actors.

Keywords: Artificial Intelligence, Startup, Healthcare Ecosystem, Clinical Processes, Value Creation, Value Destruction.

Paper type: Academic Research Paper

1 Introduction

The impact of Artificial Intelligence (AI) on healthcare systems has been widely discussed in recent literature, with particular attention to its potential to improve clinical efficiency, reduce costs, and enable more personalized care pathways (Esteva et al., 2019; Davenport and Kalakota, 2019). Beyond transforming established practices, AI has catalysed the rise of startups developing AI-based healthcare solutions (Garbuio and Lin, 2019). While previous research on AI startups has mainly focused on their business model archetypes and technological innovations, it has often overlooked their broader systemic implications. In particular, by forging novel relationships with incumbent providers, patients, regulators, and technology suppliers, AI startups enter healthcare ecosystems as new actors that may reconfigure existing structures and dynamics (Chen and Décary, 2020; Wessel et al., 2021). Depending on the solutions they provide and the ties they cultivate, these ventures may act as complementors or emerge as orchestrators capable of reshaping coordination and control (Cozzolino and Geiger, 2024). Despite their growing importance, less attention has been devoted to the role of AI startups as ecosystem actors and to their effects on clinical processes and value creation mechanisms. This study addresses this gap by investigating how healthcare AI startups contribute to the redesign of clinical processes and to the reconfiguration of value creation within healthcare ecosystems. Building on the ecosystem-as-structure perspective proposed by Adner (2017) and the role-based taxonomy of Dedehayir et al. (2018), this research conceptualizes healthcare as a complex ecosystem characterized by multilateral interdependencies among heterogeneous actors, including providers, regulators, payers, technology suppliers, and patients. Within this context, AI startups are interpreted not merely as technology vendors, but as actors that may alter activity flows, redistribute roles, and create new connections among stakeholders such as clinicians, patients, healthcare

administrators, insurance companies, and public agencies. At the same time, the effects of AI startups cannot be inferred only from their technological features or self-described value propositions. From a service and ecosystem perspective, value is not embedded in a solution itself, but emerges through use, interaction, and context (Vargo and Lusch, 2008; Grönroos and Voima, 2013). For patients, value may involve accessibility, reassurance, understanding, personalization, or continuity of care (Hennrich et al., 2024). For providers, value may involve diagnostic support, workflow efficiency, reduced burden, and improved decision-making (Garbuio and Lin, 2019). However, AI-based solutions may also generate value destruction when they introduce usability problems, workflow frictions, uncertainty, dissatisfaction, or misalignment with clinical routines (Plé and Chumpitaz Cáceres, 2010; Echeverri and Skålén, 2011). Understanding the impact of healthcare AI startups therefore requires considering both where they are positioned in the ecosystem and how their solutions are experienced by key stakeholders. The study focuses on European healthcare AI startups, a particularly relevant context because European healthcare systems are characterized by strong institutional regulation, data governance constraints, public-sector involvement, and heterogeneous national healthcare arrangements. These conditions make Europe an appropriate empirical setting for studying how AI startups enter highly regulated ecosystems and how their solutions become embedded in existing clinical and organizational routines.

Accordingly, the research question guiding this study is:

How do European healthcare AI startups modify clinical processes and create value within healthcare ecosystems?

To address this question, we first examined a sample of healthcare AI startups in order to understand what these firms actually do and where their solutions are positioned within the healthcare value chain. Rather than approaching healthcare AI startups as a homogeneous category, we analysed their value propositions, descriptions, and application domains to identify the main functions through which they contribute to the healthcare ecosystem. This preliminary mapping showed that startups tend to cluster around a limited number of recurring process domains. Consistently with prior literature on AI in healthcare, we characterized the ecosystem around two broad groups of activities: core clinical processes and support processes (Davenport and Kalakota, 2019; He et al., 2019). The first group includes startups directly involved in the production of clinical value for patients, namely Diagnosis, Treatment, and Monitoring/Follow-up. The second group includes startups that support, enable, or accelerate healthcare activities without necessarily intervening directly in patient-facing clinical encounters, namely Drug Discovery, Clinical Administrative Support, and HealthTech Infrastructure. This distinction provides the basis for analysing how different types of healthcare AI startups modify clinical processes and create value for patients, providers, and other ecosystem actors. Empirically, the study was conducted in two stages. First, a population of European healthcare AI startups was identified through Crunchbase and used to construct a functional map of the ecosystem. A selected sample of 100 startups was then analysed in depth through a hybrid natural language processing procedure combining semantic embeddings and anchor-based lexical signals. This procedure enabled the classification of each startup according to its dominant positioning across the six process categories. Second, the process-level impact of healthcare AI startups was inferred from textual comments and reviews produced by users, including both patients and care providers. This step allowed the analysis to move beyond startups' self-descriptions and examine how their solutions are perceived in practice. Reviews were cleaned, normalized, translated into English where necessary, and segmented into sentence-level units. Each sentence was then associated with one or more clinical phases — Diagnosis, Treatment, and Monitoring/Follow-up — using a hybrid approach based on keywords and semantic similarity. For each active phase, the analysis identified whether the comment expressed value creation, value destruction, or a neutral effect, and whether this value concerned patients or providers. This methodological design allows the study to connect three levels of analysis: the functional positioning of AI startups within the healthcare ecosystem, their intervention in specific clinical or support processes, and the perceived value effects experienced by key stakeholders. In this paper, preliminary findings analysing the comments related to only a sample of AI Diagnostic start-ups are shown.

A key preliminary finding is that healthcare AI startups contribute to the ecosystem through a dual functional positioning. They intervene directly in core clinical processes, especially diagnosis, treatment, and monitoring/follow-up, while also enabling clinical value creation through drug discovery, administrative support, and HealthTech infrastructure. This suggests that their role cannot be understood only through isolated technological applications but depends on the functions they perform within the healthcare value chain and on how they connect patients, providers, infrastructures, and other ecosystem actors. The review-based analysis further shows that diagnostic AI startups generate perceived value for both providers and patients, particularly by supporting clinical interpretation, decision-making, and patient understanding. However, the presence of value destruction indicates that their effects are not uniformly positive and depend on workflow integration, user experience, and alignment with clinical routines. Moreover, although the analysis focuses on diagnostic

startups, user comments suggest that their impact may extend to downstream processes such as treatment and monitoring/follow-up.

The study contributes to the literature in three ways. First, it advances an ecosystem-level understanding of healthcare AI startups, showing that their impact depends on the roles they play, the processes they address, and the relationships they reshape (Adner, 2017; Dedehayir et al., 2018; Jacobides et al., 2018). Second, it provides an empirically grounded taxonomy distinguishing between core clinical and support processes. Third, it assesses the perceived effects of AI startups in terms of value creation and value destruction, contributing to service and ecosystem research on value-in-use and value co-destruction (Vargo and Lusch, 2008; Grönroos and Voima, 2013; Plé and Chumpitaz Cáceres, 2010; Echeverri and Skålén, 2011). From a managerial and policy perspective, healthcare AI startups should be evaluated not only for technological performance, but also for workflow fit, stakeholder experience, interoperability, and their effects on adjacent clinical processes.

2 Methodology

We adopted a computationally assisted mixed-method design to link the functional positioning of healthcare AI startups with their perceived impact on clinical value creation. The analysis was conducted in two sequential stages. The first stage aimed to construct and map the functional structure of the European healthcare AI startup ecosystem. The second stage focused on assessing how AI-based solutions are perceived by users, distinguishing between value creation and value destruction across different clinical phases and stakeholder groups. The initial population was retrieved from Crunchbase and included approximately 2,500 startups headquartered in Europe and operating at the intersection of healthcare and artificial intelligence (Dalle et al., 2017). To ensure consistency with the scope of the study, we retained firms whose industry tags, descriptions, or full descriptions indicated a clear connection with healthcare, medical technologies, digital health, biotechnology, clinical software, medical devices, or related HealthTech domains. In addition, the retained firms had to show evidence of AI-related activity, including artificial intelligence, machine learning, data analytics, automation, algorithmic decision support, or comparable data-driven capabilities. Records lacking sufficient textual information to assess either healthcare relevance or AI-related activity were excluded from the in-depth classification stage. The resulting dataset contained European healthcare AI startups described through both structured and unstructured textual variables, including Organization Name, Industries, Description, and Full Description. These fields were used to infer what each startup does and where its solution is positioned within the healthcare value chain. Before the analysis, textual data were cleaned and normalized. Text was converted to lowercase, recurrent acronyms such as “EHR” were expanded, expressions such as “B2B” were harmonized, irrelevant characters were removed, and spacing was standardized. For each startup, we then constructed three textual representations: a complete representation combining industries, description, and full description; a description-based representation emphasizing the firm’s main value proposition; and an industry-only representation capturing sectoral cues. This procedure allowed us to prioritize substantive narrative information while retaining the classification value of industry labels. The functional classification relied on a hybrid natural language processing procedure combining semantic similarity and lexical anchoring (Grimmer and Stewart, 2013; Nelson, 2020). First, the textual representations of startups were transformed into vector embeddings using the BAAI/bge-base-en-v1.5 model implemented through Sentence Transformers (Reimers and Gurevych, 2019; Xiao et al., 2023). Second, the semantic proximity between each startup and each predefined functional category was calculated using cosine similarity. The classification framework included six functional categories organized into two macro-groups. The first macro-group, Clinical Process, comprised startups directly involved in clinical value creation: Diagnostics, Treatment, and Monitoring/Follow-up. The second macro-group, Support Process, comprised startups that support, enable, or accelerate healthcare activities without necessarily intervening directly in patient-facing clinical encounters: Drug Discovery, HealthTech Infrastructure, and Administrative Process (Davenport and Kalakota, 2019; He et al., 2019; Vamathevan et al., 2019). For each category, we developed a reference semantic representation, which was also transformed into an embedding. The final classification combined the embedding-based similarity score with an anchor-based lexical component based on hard, soft, and negative keyword signals. Hard signals captured highly specific terms associated with a category; soft signals captured broader but still relevant semantic cues; negative signals reduced the likelihood of misclassification when terms appeared in unrelated contexts. These components were integrated into a final score for each category, allowing each startup to be assigned to its dominant functional position (Sebastiani, 2002). In the second stage, we analysed textual user comments and reviews to infer the perceived impact of healthcare AI startups on clinical processes. This step focused on three core clinical phases: Diagnosis, Treatment, and Monitoring/Follow-up. Reviews were first cleaned and normalized, translated into English where necessary to ensure linguistic comparability, and segmented into sentence-level units. Each sentence was then

classified according to the clinical phase or phases to which it referred. Phase identification relied on a hybrid procedure combining phase-specific keywords with semantic similarity based on embeddings. Since a sentence could refer to more than one clinical activity, the model allowed multiple phases to be activated simultaneously.

For each activated clinical phase, we assessed whether the sentence expressed value creation, value destruction, or a neutral effect. We define value creation as the perceived improvement in the actor's ability to accomplish a relevant clinical or care-related task, consistent with service-dominant and service-logic perspectives that conceptualize value as value-in-use and value-in-context rather than as an intrinsic property of the provider's offering (Vargo and Lusch, 2008; Grönroos and Voima, 2013). As a result, value creation captured perceived benefits such as improved efficiency, better clinical support, smoother processes, enhanced user experience, or increased accessibility. Conversely, we define value destruction as the negative counterpart of value creation, that is, as a perceived deterioration in an actor's ability to accomplish a relevant clinical or care-related task. This interpretation follows Schulz et al. (2021), who conceptualize value co-creation and value co-destruction as "two sides of the same coin", with value co-destruction understood as the antonym of value co-creation. This approach is consistent with earlier work defining value co-destruction as an interactional process that results in a decline in at least one actor's well-being (Plé and Chumpitaz Cáceres, 2010), as well as later reviews that frame value co-creation and co-destruction as alternative outcomes of interactive value formation (Echeverri and Skålén, 2021). As a result, value destruction captured negative effects such as problems, delays, inefficiencies, dissatisfaction, usability issues, or workflow frictions. The analysis also identified the stakeholder to whom the value effect referred, distinguishing between providers and patients. As in the functional classification stage, this assessment combined lexical signals and semantic embeddings. When a sentence activated a clinical phase but did not contain clear signals of either value creation or value destruction, it was classified as neutral.

3 Results

The initial query on Crunchbase returned 2.499 European startups active in the field of Healthcare AI. A first look at their geographical distribution (Figs. 1–2) reveals a strong concentration in high-GDP countries: more than 27% of firms are headquartered in the United Kingdom (676 start-ups, with 437 only in London), followed—at a considerable distance—by Germany (273 start-ups, 10.9%) and France (250 firms, 10.0%). Together, these three countries account for almost 48% of all startups. The remaining landscape is distributed across Spain and Switzerland, both with 155 firms (6.2% each), Italy and the Netherlands, both with 118 firms (4.7% each), and a long tail of smaller national ecosystems. This captures a core Western European concentration, although the broader dataset also highlights a more visible presence of Mediterranean, Nordic, Central and Eastern European countries.

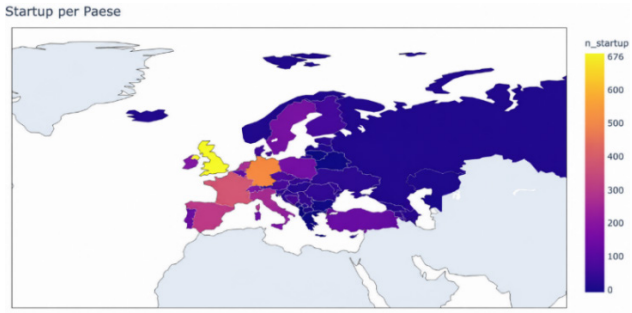


Figure 1: Distribution of European AI startups (Heat map) - Source: authors' elaboration based on Crunchbase data.

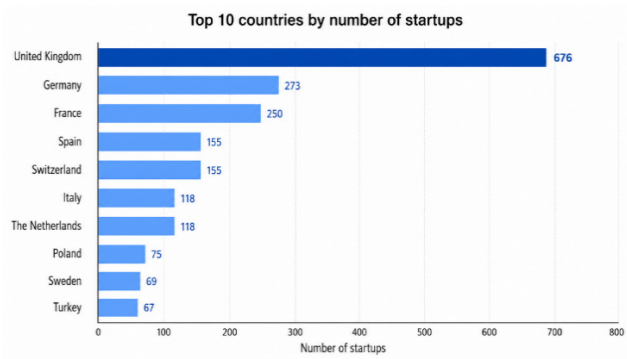


Figure 2: Distribution of European AI startups - Source: authors' elaboration based on Crunchbase data.

As for the founding date, available for 2,338 startups, the median founding year is 2019, and around 46% of firms with available data were founded between 2020 and 2026. This reinforces the idea that healthcare AI in Europe is a young and rapidly expanding entrepreneurial field. From an organizational perspective, the sample is dominated by small ventures. Out of 2,399 startups with available employee data, 1,321 firms (55.1%) employ between 1 and 10 people, while another 850 firms (35.4%) fall within the 11–50 employees range. Overall, more than 90% of firms with available data have fewer than 50 employees. Only 39 startups exceed 250 employees, equal to 1.6% of the employee-available sample. This confirms that the European healthcare AI landscape is mostly composed of early-stage or small-scale ventures rather than large consolidated companies.

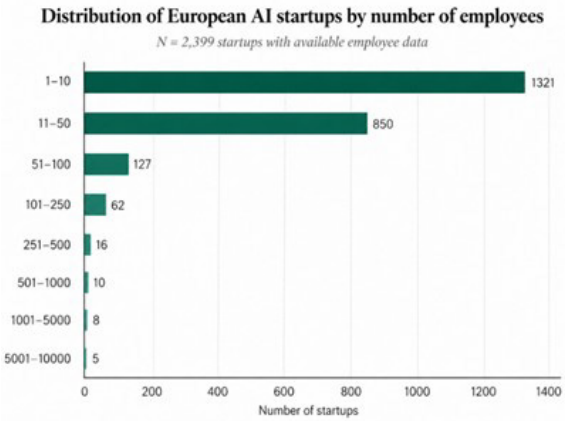


Figure 3: Distribution of European AI startups by number of employees – Source: authors' elaboration based on Crunchbase data.

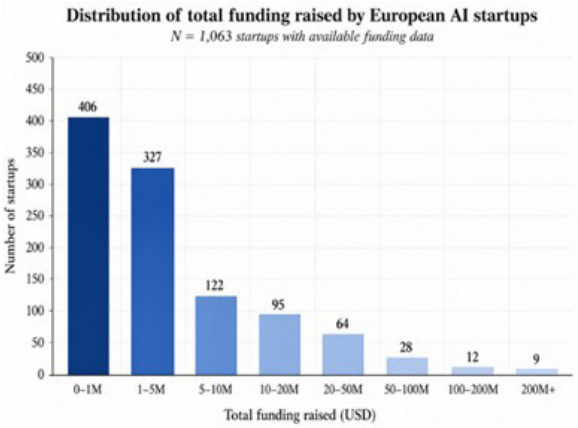


Figure 4: Distribution of total funding raised by European startups – Source: authors' elaboration based on Crunchbase data.

Given the heterogeneity of the startups identified, revenue remains a useful proxy for economic maturity and market traction. In the dataset, 1,219 records report an estimated revenue range. The majority of these firms, 706 startups (57.9%), fall within the $1M–$10M revenue bracket, while 412 firms (33.8%) report revenues below $1M. Only 80 firms (6.6%) are positioned in the $10M–$50M range, and very few exceed this threshold. This suggests that most healthcare AI startups have reached some degree of monetization, but remain far from the revenue scale typically associated with mature medium-sized enterprises. Total funding raised shows a highly skewed distribution. Funding data are available for 1,063 startups, which together account for approximately $12.83 billion in total funding. However, the median funding amount is only around $1.95 million, while the average is approximately $12.1 million, indicating the presence of a small number of highly funded outliers. Almost 89.4% of funded startups have raised less than $20 million, while only 21 firms — around 2.0% of those with funding data — have raised more than $100 million. The largest funded companies include Babylon, Isomorphic Labs, Cera, Exscientia, Dental Monitoring, Neko Health, Huma, BenevolentAI, Flo Health and Ada Health. The concentration is particularly strong: the top 20 funded companies account for approximately 44.6% of all funding captured in the dataset. The investor landscape also confirms the central role of public programs, accelerators and early-stage funding schemes. Based on the "Top 5 Investors" field, Innovate UK is the most recurrent investor, appearing in 53 funded companies, followed by EISMEA with 40, Bpifrance and the European Innovation Council with 33 each, Venture Kick with 32, and Google for Startups with 30. Other recurrent investors include Techstars, Plug and Play, Antler, Entrepreneurs First, EIT Health and MassChallenge. The prevalence of public innovation agencies, European funding bodies and accelerator programs suggests that institutional and semi-institutional support remains a structural component of the European healthcare AI ecosystem.

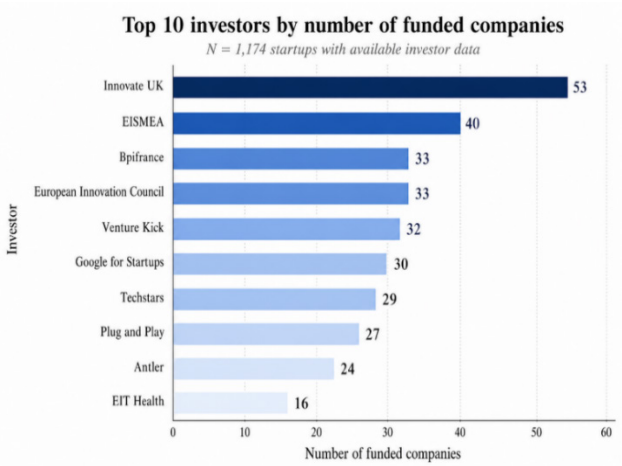


Figure 5: Top 10 investors by number of funded companies
Source: authors' elaboration based on Crunchbase data.

Among firms with last-funding-date data, almost 70% received their latest funding round in 2022 or later, and around 44% in 2024 or later. This suggests that the sector has remained financially active in the post-pandemic period, despite broader venture-capital selectivity. Furthermore, among startups with a recorded last funding type, Seed, Pre-Seed and Angel rounds account for about 50.7% of cases. By contrast, Series A and Series B together represent only around 11.1%. This supports the interpretation of a pipeline rich in early-stage experimentation but thinner in later-stage scale-up. Overall, this screening confirms several preliminary findings. First, the United Kingdom emerges as the most vital hub and the main entry point to the European market, followed by Germany and France, while southern and eastern Europe remain comparatively underrepresented. Second, the majority of startups are still in an "early scale" phase, with small organizational size, revenues concentrated in the lower brackets, and limited access to large-scale funding. Third, the ecosystem is strongly shaped by public programs and corporate accelerators, which play a central role in sustaining the pipeline of healthcare AI ventures. These insights guided the subsequent selection of the sample for in-depth analysis.

3.1 Preliminary findings

The preliminary analysis suggests that European healthcare AI startups are distributed across both core clinical processes and support processes, indicating that their role in healthcare ecosystems is not limited to direct patient-facing clinical applications. Instead, these firms contribute to healthcare transformation through a combination of clinical, infrastructural, administrative, and drug discovery-oriented activities. First, the functional classification of the startups shows a relatively balanced distribution between the two macro-groups. Core clinical processes account for 53.6% of the sample, while support processes account for 46.4%. This indicates that healthcare AI startups are only slightly more concentrated in directly clinical domains than in enabling or infrastructural domains. The result suggests that AI-driven transformation in healthcare does not occur exclusively at the bedside or within clinical decision-making, but also through the reconfiguration of the supporting conditions that make clinical work possible. At the level of specific functions, the two most represented categories are Drug Discovery and Diagnostics, which account for 27.8% and 27.1% of the sample respectively. This finding indicates that healthcare AI startups are particularly active in domains where AI can process complex data, identify patterns, and support knowledge-intensive tasks. Diagnostics represents the main clinical entry point for AI startups, while Drug Discovery represents the strongest support-oriented domain. By contrast, Treatment accounts for only 12.0% of the sample and Administrative Process for 3.0%, suggesting that some areas of the healthcare value chain remain less populated by AI startups or may present higher barriers to entry.

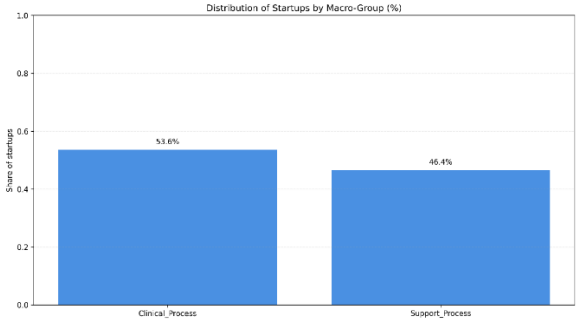


Figure 6: Distribution of Startups by Macro-Group

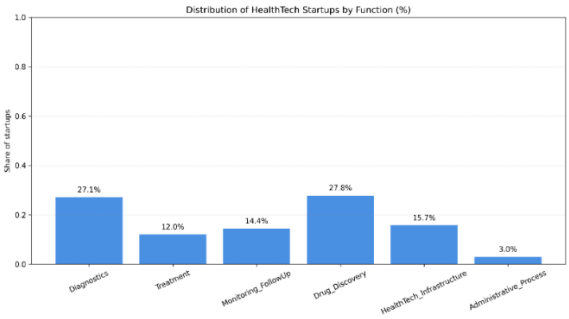


Figure 7: Distribution of HealthTech Startups by Function

The radial classification map (Figure 8) further suggests that the ecosystem is not composed only of sharply specialized firms. Several startups occupy intermediate positions between functional poles, indicating the presence of hybrid or polyfunctional business models. This is particularly relevant from an ecosystem perspective because it shows that AI startups may simultaneously contribute to multiple phases of the healthcare process, for instance by combining diagnostic support with monitoring capabilities, or infrastructure provision with clinical analytics. This hybridity points to a more complex form of ecosystem participation, where startups cannot always be reduced to a single clinical or support function.

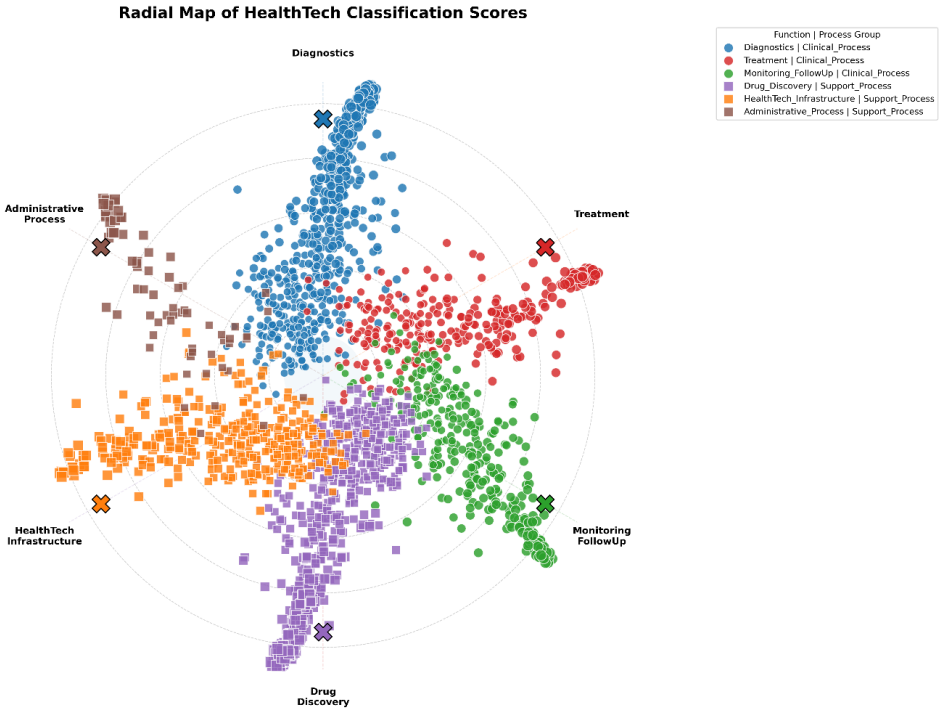


Figure 8: radial classification map of the AI-start-ups within clusters

Second, the analysis of user comments and reviews provides an initial indication of how these startups are perceived to affect clinical processes in practice. The review-based analysis focused preliminary on the Diagnostic phase and distinguished between value creation, value destruction, and neutral effects for two stakeholder groups: providers and patients. For providers, diagnosis contains 156 classified events, with a net value score of +18; for patients, it contains 140 events, with a net value score of +10. This suggests that diagnostic AI solutions generate visible value for both stakeholder groups, particularly by supporting clinical interpretation, decision-making, or patient understanding. However, the presence of both positive and negative events also indicates that diagnostic AI is not experienced as uniformly beneficial (see Figure 9 and 10). Its value appears to depend on how well the solution is integrated into clinical workflows, how reliable users perceive it to be, and whether it improves rather than complicates the diagnostic process.

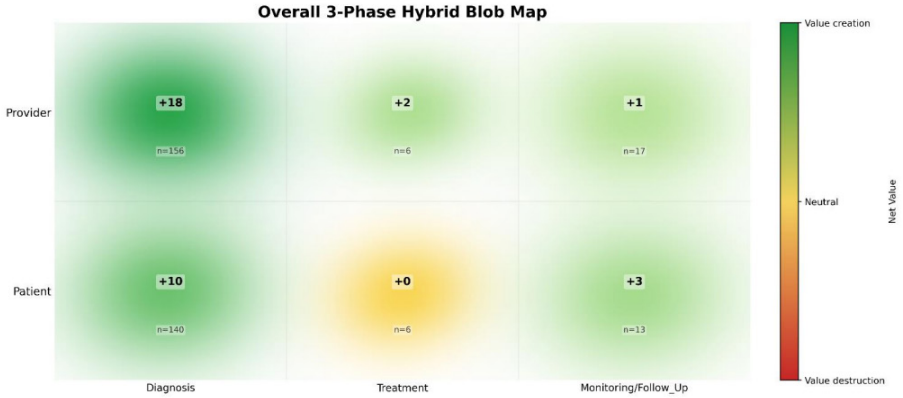


Figure 9: Value formation Heatmap for the diagnostic category sample

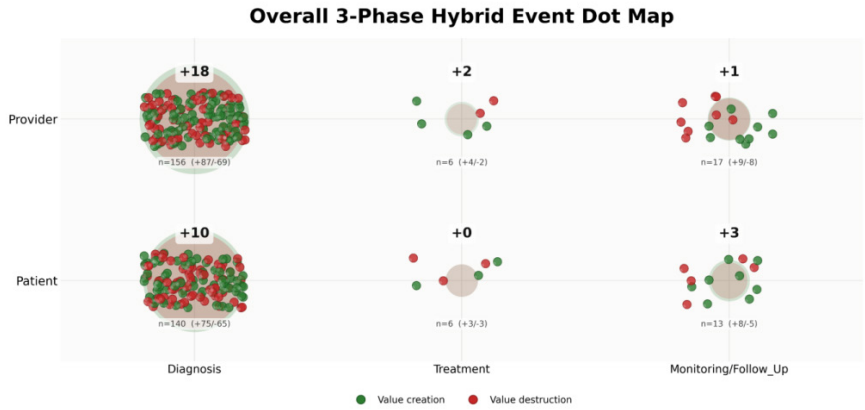


Figure 10: Value formation Dotmap for the diagnostic category sample

The review-based value analysis was conducted on diagnostic startups, but the comments were coded to capture whether the perceived value of diagnostic solutions extended beyond diagnosis and affected other clinical phases, such as treatment and monitoring/follow-up. Although these startups are primarily positioned in the diagnostic phase, users' comments sometimes reveal perceived effects that extend beyond diagnosis and spill over into downstream clinical processes. Treatment records six provider-related and six patient-related events, with limited net value effects, while Monitoring/Follow-up shows a somewhat stronger but still modest presence, with 17 provider-related and 13 patient-related events. These results suggest that diagnostic AI startups are perceived to create value mainly in relation to diagnostic activities, but their impact may also reverberate across the broader clinical pathway, especially when diagnostic outputs influence therapeutic decisions, patient management, or follow-up practices.

4 Conclusions

This study contributes to the literature in three main ways. First, it advances an ecosystem-level understanding of healthcare AI startups. Prior research has widely discussed the potential of AI to improve healthcare delivery, for instance by enhancing diagnostic accuracy, clinical efficiency, personalization, and decision-making (Esteva et al., 2019; Davenport and Kalakota, 2019). However, much of this literature remains centred on the technological capabilities of AI systems or on the business models of AI-based ventures (Garbuio and Lin, 2019). By contrast, this study shifts attention from the technology itself to the ecosystem role played by AI startups. Building on the ecosystem-as-structure perspective (Adner, 2017) and on ecosystem actor-role approaches (Dedehayir et al., 2018; Jacobides et al., 2018), we show that the impact of healthcare AI startups depends not only on what the technology can do, but also on where the startup is positioned in the healthcare activity system, which actors it connects, and which interdependencies it modifies. In this sense, AI startups are not interpreted merely as providers of technological solutions, but as ecosystem actors that may reconfigure relationships among patients, providers, infrastructures, regulators, and other healthcare stakeholders. This contributes to a less technology-deterministic understanding of AI in healthcare, showing that the transformative potential of AI emerges through the roles, relationships, and coordination mechanisms that

startups activate within the ecosystem. Second, the study provides an empirically grounded functional taxonomy of healthcare AI startups. Our analysis offers a more granular classification than previous studies (e.g., Garbuio and Lin) based on the actual function performed by startups within the healthcare value chain. Specifically, we distinguish between startups operating in core clinical processes — Diagnosis, Treatment, and Monitoring/Follow-up — and startups operating in support processes — Drug Discovery, Clinical Administrative Support, and HealthTech Infrastructure. This taxonomy is original because it is not derived only from theoretical categories or industry labels, but from the systematic analysis of a large number of startup descriptions, value propositions, and application domains. It therefore provides a structured map of how healthcare AI startups populate different areas of the ecosystem and clarifies that AI-driven transformation occurs both directly, through patient-facing clinical processes, and indirectly, through the infrastructures, administrative systems, and drug discovery activities that enable clinical value creation. These start-ups intervene directly in core clinical processes, especially diagnosis, where their effects are most visible in user reviews. Furthermore, they play a substantial enabling role through support processes such as Drug Discovery, Clinical Administrative Support, and HealthTech Infrastructure. Third, research on AI in healthcare often assumes that AI adoption produces positive effects, such as efficiency gains, improved decision-making, or better patient outcomes (Davenport and Kalakota, 2019; Topol, 2019). However, service and ecosystem research suggests that value is not embedded in a solution itself, but emerges through use, interaction, and context (Vargo and Lusch, 2008; Grönroos and Voima, 2013). By analysing user comments from patients and providers, this study moves beyond startups' self-representations and assesses the perceived effects of AI solutions in practice. This allows us to show that healthcare AI startups may create value but may also generate value destruction. The originality of this contribution lies hence in connecting the functional positioning of AI startups with their stakeholder-specific value effects, showing that the impact of healthcare AI is systemic, relational, and contingent on how solutions are embedded in clinical routines. From a managerial perspective, the evidence from diagnostic startups highlights that Diagnostic AI solutions appear to generate visible value for both providers and patients, but their impact may also extend to treatment decisions and follow-up practices. Managers should therefore assess diagnostic AI not as a narrow diagnostic tool, but as a technology that may affect downstream clinical processes and interdependencies among actors. Accordingly, healthcare managers should evaluate and implement diagnostic AI through a process-integration lens, considering not only accuracy and technical performance, but also workflow fit, stakeholder experience, data interoperability, and the potential effects on subsequent clinical decisions and care coordination.

5 Ethics Declaration

This research relied on publicly available secondary data and did not involve direct interaction with human participants.

6 AI tools Declaration

AI tools were used to support language editing and text refinement. The authors retained full responsibility for the conceptual development, analysis, interpretation of results, and final content of the paper.

References

Adner, R. (2017) Ecosystem as structure: An actionable construct for strategy, *Journal of Management*, Vol. 43, No. 1, pp. 39–58. https://doi.org/10.1177/0149206316678451.

Chen, M. and Décary, M. (2020) Artificial intelligence in healthcare: An essential guide for health leaders, *Healthcare Management Forum*, Vol. 33, No. 1, pp. 10–18. https://doi.org/10.1177/0840470419873123.

Cozzolino, A. and Geiger, D. (2024) Ecosystem disruption and regulatory positioning: Entry strategies of digital health startup orchestrators and complementors, *Research Policy*, Vol. 53, No. 10, 104913. https://doi.org/10.1016/j.respol.2024.104913.

Dalle, J.-M., den Besten, M., & Menon, C. (2017). *Using Crunchbase for economic and managerial research*. OECD Science, Technology and Industry Working Papers, No. 2017/08. OECD Publishing.

Davenport, T. and Kalakota, R. (2019) The potential for artificial intelligence in healthcare, *Future Healthcare Journal*, Vol. 6, No. 2, pp. 94–98. https://doi.org/10.7861/futurehosp.6-2-94.

Dedehayir, O., Mäkinen, S. J. and Ortt, J. R. (2018) Roles during innovation ecosystem genesis: A literature review, *Technological Forecasting and Social Change*, Vol. 136, pp. 18-29.

Echeverri, P., & Skålén, P. (2011). Co-creation and co-destruction: A practice-theory based study of interactive value formation. *Marketing Theory*, *11*(3), 351–373.

Esteva, A., Robicquet, A., Ramsundar, B., Kuleshov, V., DePristo, M., Chou, K., Cui, C., Corrado, G., Thrun, S. and Dean, J. (2019) A guide to deep learning in healthcare, *Nature Medicine*, Vol. 25, No. 1, pp. 24–29. https://doi.org/10.1038/s41591-018-0316-z.

Garbuio, M. and Lin, N. (2019) Artificial intelligence as a growth engine for health care startups: Emerging business models, *California Management Review*, Vol. 61, No. 2, pp. 59-83.

Greaves, F., Ramirez-Cano, D., Millett, C., Darzi, A., & Donaldson, L. (2013). Use of sentiment analysis for capturing patient experience from free-text comments posted online. *Journal of Medical Internet Research*, *15*(11), e239.

Grimmer, J., & Stewart, B. M. (2013). Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political Analysis*, *21*(3), 267–297.

Grönroos, C., & Voima, P. (2013). Critical service logic: Making sense of value creation and co-creation. *Journal of the Academy of Marketing Science*, *41*, 133–150.

He, J., Baxter, S. L., Xu, J., Xu, J., Zhou, X., & Zhang, K. (2019). The practical implementation of artificial intelligence technologies in medicine. *Nature Medicine*, *25*, 30–36.

Jacobides, M. G., Cennamo, C. and Gawer, A. (2018) Towards a theory of ecosystems, *Strategic Management Journal*, Vol. 39, No. 8, pp. 2255–2276. https://doi.org/10.1002/smj.2904.

Nelson, L. K. (2020). Computational grounded theory: A methodological framework. *Sociological Methods & Research*, *49*(1), 3–42.

Plé, L., & Chumpitaz Cáceres, R. (2010). Not always co-creation: Introducing interactional co-destruction of value in service-dominant logic. *Journal of Services Marketing*, *24*(6), 430–437.

Reimers, N., & Gurevych, I. (2019). Sentence-BERT: Sentence embeddings using Siamese BERT-networks. In *Proceedings of EMNLP-IJCNLP 2019*.

Sebastiani, F. (2002). Machine learning in automated text categorization. *ACM Computing Surveys*, *34*(1), 1–47.

Vamathevan, J., Clark, D., Czodrowski, P., Dunham, I., Ferran, E., Lee, G., Li, B., Madabhushi, A., Shah, P., Spitzer, M., & Zhao, S. (2019). Applications of machine learning in drug discovery and development. *Nature Reviews Drug Discovery*, *18*, 463–477.

Vargo, S. L., & Lusch, R. F. (2008). Service-dominant logic: Continuing the evolution. *Journal of the Academy of Marketing Science*, *36*(1), 1–10.

Wessel, L., Baiyere, A., Ologeanu-Taddei, R., Cha, J. and Blegind Jensen, T. (2021) Unpacking the difference between digital transformation and IT-enabled organizational transformation, *Journal of the Association for Information Systems*, Vol. 22, No. 1, pp. 102–129. https://doi.org/10.17705/1jais.00655.

Xiao, S., Liu, Z., Zhang, P., & Muennighoff, N. (2023). *C-Pack: Packaged resources to advance general Chinese embedding*. arXiv preprint.

Organizational Capabilities for Knowledge Ecosystems in Public Healthcare: Towards Intelligent and Sustainable Transformation

Concetta Lucia Cristofaro[1], Marzia Ventura[2], Anna Maria Melina[3], Walter Vesperi[4]

[1] University eCampus, Novedrate, Italy
[2] University of Catanzaro, Catanzaro, Italy
[3] University of Messina, Messina, Italy
[4] University of Palermo, Palermo, Italy

Abstract

In the current context, public health systems are facing an unprecedented challenge, namely, being able to manage increasing complexity resulting from demographic aging, chronicity and the need to integrate frontier technologies in contexts of limited resources. In this scenario, the healthcare organization evolves from a simple bureaucratic structure to a "knowledge ecosystem". the healthcare organization becomes a dynamic network in which the value lies in individual services, but also and above all in the ability to generate, share and apply multidimensional knowledge. This study was born with the aim of exploring how organizational capabilities act as a fundamental pillar to support and, above all, make these ecosystems sustainable in the long term. The central objective of the research is to identify the key organizational capabilities necessary to overcome information fragmentation and promote real integration between the different players in the system (clinicians, managers, patients and technological partners). To achieve this aim, the study adopts a qualitative methodology based on interviews with strategic decision-makers, highlighting three macro-categories of essential capabilities.
First, the absorptive capacity emerges, therefore, the organization's propensity and ability to identify, assimilate and apply all those external innovations and complex clinical data to translate them into efficient operational protocols. Secondly, relational and collaborative skills, essential for breaking down functional silos and promoting more widespread leadership models. These capabilities allow the transition from hierarchical personnel management to a community of practice that enhances human capital as the driver of the ecosystem. Finally, strategic digital capabilities, understood not only as infrastructural equipment, but as the organization's ability to manage data interoperability and mitigate phenomena of resistance to change and technostress. The results suggest that the sustainability of a knowledge ecosystem does not depend exclusively on technological investment, but also on a strategic alignment between organizational culture and knowledge management processes. Without a solid foundation of organizational capabilities, digital infrastructures risk becoming mere repositories of unused data. The research contributes to the debate on healthcare management by proposing a model in which performance is directly proportional to the company's ability to continuously regenerate its knowledge assets.

Keywords: knowledge eco-system, public health care, organizational capabilities, strategic choice, interview

Paper type: Academic Research Paper

1 Introduction

Socio-demographic tendencies, rapid technological innovation and institutional reforms are some of the pressures for change driving contemporary public healthcare systems to radically evolve. At the same time, healthcare organizations must respond to growing and evolving social expectations, while constrained by limited financial resources (Khanuja, et al., 2025; Ahmad, et al.,2022; Binci, et al., 2022). These conditions are challenging traditional organizational models, exposing their limitations and consequential in reduced performance. Changes in healthcare organizations and processes have a direct impact on well-being and quality of life (Öberg and Lundberg, 2022; Maracine and Scarlat, 2009).

To address these limitations, healthcare organizations must reorient their strategic choices, and internal organizational process, by focusing on the enabling role of knowledge and organizational capabilities from a new strategic perspective (Khanuja, et al., 2025). In fact, healthcare organizations must support and implement the creation of knowledge ecosystems. From this perspective, this study aims to analyze knowledge ecosystems in the context of public healthcare. Therefore, this paper, from the theoretical perspective of knowledge management studies, aims to provide an overview of the elements and organizational capabilities that shape knowledge ecosystems. The study adopts an exploratory-interpretive qualitative methodology, with a process-oriented approach to organizational change. This methodological approach is particularly suited to capturing the dynamic, relational, and multilevel nature of knowledge ecosystems within complex public systems.

Empirical data was collected through in-depth, semi-structured interviews with key stakeholders involved in the governance and implementation of healthcare reform. Data analysis was conducted using thematic analysis, with the aim of identifying recurring patterns related to the processes of knowledge production, sharing, and valorization within processes. Emphasis was paid to knowledge flows between organizations, as well as the role of digital and collaborative technologies in supporting these dynamics.

This study offers a theoretical contribution to the reconfiguration of public healthcare systems as knowledge ecosystems, enriching the debate on knowledge management and organizational capacity. Overall, the findings of this study highlight that the sustainability and resilience of public healthcare systems depend on the ability to cultivate knowledge ecosystems capable of integrating human, organizational, and institutional intelligence.

From this perspective, the paper proposes an ecosystemic view of strategic change, in which capabilities are not isolated properties of individual organizations, but emerge from the interaction between institutional, organizational, and community actors. Knowledge ecosystems are described as dynamic configurations traversed by structural tensions: between standardization and local adaptation, between control and professional autonomy, between digital rationalization and human judgment.

The findings of this study have significant practical implications for public policymakers, managers, and professionals engaged in the design and governance of smart and sustainable organizations and communities.

2 Theoretical Background

2.1 Knowledge ecosystems in healthcare

Previous studies have analyzed knowledge ecosystems in various contexts, including regional startups and internal organizational environments, highlighting their structural characteristics and the composition of their actors (Clarysse et al., 2014; Miller et al., 2016; Sharma & Bhattacharya, 2013). However, much research has focused on the transition from knowledge ecosystems to business ecosystems, while limited attention has been paid to the organizational modalities required to achieve knowledge development objectives (Cobben et al., 2022).

In healthcare, the literature highlights how their development follows similar trajectories across countries, reflecting the broader evolution of organizational theories. In this context, Kernick (2004) identifies four phases in the development of the National Health Service (NHS) that help explain the current relevance of the ecosystem approach. The first three phases are characterized by a progressive transition from command-and-control models, based on rigid hierarchies and the mechanistic metaphor of organization, to market-oriented configurations and subsequently to hybrid forms that integrate cooperation and competition between public and private actors. The fourth phase introduces a radically different vision, interpreting the healthcare system as a set of interrelated systems that interact in a nonlinear fashion, in which the unit of analysis shifts from the individual patient to the patient-professional relationship. From this perspective, healthcare systems are conceived as complex ecosystems, characterized by diversity, connectivity, feedback, and self-organizing rules that foster the emergence of new behavioral patterns (Kernick, 2004).

The ecosystem approach emphasizes the impossibility of understanding individual actors in isolation, favoring instead a systemic and relational analysis (Royston & Dick, 1998). Broadly speaking, a healthcare ecosystem can be defined as a complex system in which a multitude of actors are intertwined – patients, healthcare professionals, healthcare facilities, public administrations and regulatory bodies – and a variety of resources, such as professional skills, clinical knowledge, technologies, tools and regulatory devices (Di Nauta et al., 2017). These elements do not operate separately, but as highlighted by Frow et al., (2016), they are interconnected through value propositions, which, within a complex framework, configure a network of relationships. Actors are motivated to actively participate in the ecosystem to the extent that sharing resources allows access to benefits that include and exceed the results that an individual action could bring. In this context, the healthcare ecosystem can be defined as a dynamic system; In the interaction of actors, a continuous reallocation and recombination of resources over time brings variance to the attractiveness of value propositions; these, however, through their configuration, define the modalities of interaction. From this perspective, practices play a primary role; that is, as sets of activities and interactions that are constructed in social contexts and that actors continually modify based on how they interpret the sociality of the construction (Kjellberg & Helgesson, 2007). The ecosystem evolves based on the value of the activities selected and modified by the actors involved. Co-creators of value possess resources that, when integrated with others, will become a benefit. The sharing of a patient's clinical data between a hospital and others (such as specialists or healthcare providers) enables personalized treatment. The importance of interactional collaboration in the healthcare system has made co-creation processes particularly attractive. Healthcare, traditionally unidirectional and

passive, where healthcare providers deliver care and patients receive it, now places patients in an active position. Recently, a new perspective has emerged that recognizes patients' active and participatory role in managing their health. This approach, known as patient-centered, focuses on designing healthcare services based on patients' specific needs, preferences, and conditions, leading to improved clinical outcomes and greater efficiency of the system as a whole. Adopting a patient-centered approach has expanded the number and variety of stakeholders involved in co-creating value. In addition to healthcare professionals, family and friends are increasingly being included, offering informational, emotional, and practical support throughout the care journey. Furthermore, there has been an increase in collaborative activities, which are no longer limited to sharing clinical decisions among experts but also include initiatives to improve patients' psychological and emotional well-being and the quality of their care experience. Despite these developments, healthcare literature has long neglected a systematic analysis of the interactions between the various actors involved and the collaborative practices that characterize this ecosystem. Only recently, with contributions such as that of Frow et al. (2016), has it become clearer how important it is to study relational dynamics and the mechanisms of value co-creation in healthcare from an ecosystemic perspective, capable of capturing the complexity, dynamism, and interdependence of this field.

From the above, it emerges that the concept of knowledge ecosystems has gained increasing over the years, as organizations increasingly move toward models that appear more interconnected and collaborative. A knowledge ecosystem is a dynamic network of actors, institutions, and technologies engaged in the co-creation, exchange, and application of knowledge. In the healthcare sector, knowledge ecosystems are particularly relevant due to the complexity of the sector itself, which involves multiple stakeholders, including physicians, administrators, policymakers, patients, and technology providers. These ecosystems are therefore characterized by distributed knowledge sources, interorganizational collaboration, and, above all, ongoing learning and adaptation processes.

Therefore, knowledge ecosystems in healthcare differ from traditional organizational models in that they emphasize relationships and flows rather than structures and hierarchies. Knowledge is not confined within organizational boundaries, but circulates through networks. All of this enables innovation and continuous improvement in service delivery.

2.2 Organizational capabilities in knowledge ecosystem

Hetemi et al (2026) argue that knowledge ecosystems differ from traditional hierarchical organizations in their organizational specificities (Kretschmer et al., 2022) namely, the nature of authority and control mechanisms; motivational structures and incentives for participation; and finally governance and coordination practices. Healthcare organizations are no longer viewed merely as service providers but as complex, knowledge-intensive systems. The growing relevance of knowledge as a strategic resource has led to the emergence of the concept of knowledge ecosystems**,** where multiple actors interact to produce, share, and apply knowledge across organizational and institutional boundaries. This transformation requires a shift in perspective: from hierarchical and siloed structures to more collaborative, adaptive, and networked configurations. From this perspective, knowledge ecosystems can be viewed as a community based on individuals and organizations that continuously interact and primarily co-evolve in their roles and capabilities to support the growth and sustainability of the ecosystems themselves. Specifically, as highlighted by Abbate et al. (2021), both individual and organizational capabilities can contribute to the creation of new knowledgeIn healthcare systems, organizational capabilities are important to understanding how can evolve toward knowledge ecosystems. As emphasized (Amit and Schoemaker, 1993, Teece et al., 1997), organizational capabilities are defined as a otganization's ability to employ its resources, whether tangible or intangible, to perform a task or activity in order to improve performance. Helfat and Peteraf (2003) define organizational capability as "an organization's ability to perform a coordinated set of tasks, using organizational resources, to achieve a particular end result." (Dosi et al., 2000) argue that organizational capabilities are critical to firms' ability to effectively solve their organizational problems. Researchers distinguish between different organizational capabilities. Collis (1994) proposed four categories of organizational capabilities. The first "concerns the ability to perform the basic functional activities of the organization" (Collis, 1994). The second category concerns dynamic improvements to the organization's activities, such as continuous improvement activities. The third category is "recognizing the intrinsic value of other resources or developing new strategies before competitors" (Collis, 1994). The fourth category is labeled "higher-order" or "meta-capabilities" and refers to learning capabilities. Winter (2003) proposes that there are zero-level capabilities, also called operational or ordinary capabilities, which he defines as those that allow the firm to earn a living in the present. He then explains that there are first-order capabilities that modify and transform zero-level capabilities. He also

suggests, similarly to Collis (1994), that there are higher-order capabilities that operate on top of first-level capabilities. Based on capabilities, organizations must be able to identify opportunities and threats, seize opportunities through various strategic actions and reconfigure resources to adapt to increasingly evolving and changing contexts. In the healthcare context, these capabilities are particularly important because of high levels of regulation, strong professional autonomy, and, above all, complex decision-making processes. Capabilities in the healthcare sector are therefore not limited to technical skills but also include managerial, relational and institutional dimensions. Organizational skills are considered mechanisms that enable (or hinder) the flow of knowledge between stakeholders, institutional levels, and disciplinary fields. A literature analysis shows that three main categories of capabilities are considered important. The first group includes relational capabilities, such as the ability to build trust and collaboration, as well as network coordination. a second group includes digital capabilities, therefore, data literacy and above all the use of digital tools, finally the third group contains institutional capabilities, which include alignment with policies and regulations and governance mechanisms.

3 Methodology

This study adopts a qualitative, exploratory research design (Yin, 2015), suitable for investigating complex and evolving phenomena such as knowledge ecosystems. The purpose of this paper is to explore how organizational capabilities support the development of knowledge ecosystems in public healthcare. Specifically, the study aims to identify the key capabilities required to sustain knowledge ecosystems. The study adopts semistructured interviews to identify critical factors in the knowledge ecosystem of the Calabrian healthcare system. This approach is consistent with an exploratory and interpretative design already applied to the study of public healthcare ecosystems. The interviews were conducted with five public policy makers operating in the healthcare sector of the Calabrian regional healthcare system, selected based on their strategic role in the system's decision-making and governance processes. This focus on decision makers is consistent with studies that analyze knowledge ecosystems starting with key governance and coordination actors (Primario et al., 2024; Eisenhardt and Graebner, 2007). Semistructured interviews were conducted, with an average duration of approximately 45 minutes each. Each interview was conducted in person by two researchers. One served as the primary interviewer and one as an observer, responsible for taking notes on the context, nonverbal cues, and interactive dynamics. The interviews were conducted with the objective to identify and validate the organizational capabilities of the healthcare knowledge ecosystem, identified in the literature.

4 Results

This study provides a comprehensive framework for understanding knowledge ecosystems within the public health sector by integrating three pivotal theoretical strands: knowledge management, innovation ecosystems, and interorganizational relationships. Despite the strategic importance of this convergence, identifying and defining organizational capabilities remains a significant challenge due to terminological ambiguity. To address this gap, the research conceptualizes organizational capabilities as enabling (or constraining) mechanisms that govern the multidirectional flow of knowledge across diverse actors, institutional echelons, and disciplinary domains.

Table 1 summarises some organizational capabilities emerging from the literature analysis and interviews.

Table 1 - Organizational capabilities in knowledge ecosystem

	Organizational capabilities	Description	Main References
1	Leadership and strategic vision	Shared vision, sponsorship, learning orientation	Al Amiri et al., 2020; Pellegrini et al., 2020; Spanos et al., 2024; Doherty et al., 2021
2	Governance structures	Roles, responsibilities, inter-institutional coordination	Mondal et al., 2025; Enticott et al., 2020; Rabelo Neto et al., 2024, Vesperi, et al., 2021
3	Organizational culture and trust	Openness, sharing, psychological safety	Cormican et al., 2021; Kavalić et al., 2021; Preciado-Ortiz & Santander-Salmon, 2024; Secundo, et al., 2019
4	Staff Skills and abilities	Staff skillsand capabilities	Laukka et al., 2022; Spanos et al., 2024; Kavalić et al., 2021, Tessema, 2025, Secundo, et al., 2019
5	Information systems and digital infrastructure	Interoperability, quality and access to data	Laukka et al., 2022; Enticott et al., 2020; Wahid et al., 2024, Tessema, 2025

	Organizational capabilities	Description	Main References
6	Knowledge management processes	Acquisition, transfer, share	Mardian et al., 2025; Al Amiri et al., 2023; Firmansyah, 2024, Tessema, 2025
7	Coordination and collaboration	Networks, partnerships, communities of practice	Leal et al., 2025; Rabelo Neto et al., 2024; Cormican et al., 2021; Vesperi, et al., 2021
8	Policies and regulations	Consistency with national, local and corporatehealth regulations, rules and plans	Mondal et al., 2025; Wahid et al., 2024; Doherty et al., 2021; Vesperi, et al., 2021
9	Incentives and evaluation systems	Awards, career, collaboration-oriented indicators	Cormican et al., 2021; Kavalić et al., 2021; Cormican meta-analisi cit. Nguyen et al., 2021
10	Change Management	Change support, communication, training	Spanos et al., 2024; Doherty et al., 2021; Firmansyah, 2024
11	Stakeholder engagement	Clinicians, patients, communities, universities	Enticott et al., 2020; Leal et al., 2025; Doherty et al., 2021, Tessema, D. B. (2025)
12	Resources and economic sustainability	Budget, time, dedicated staff	Kavalić et al., 2021; Al Amiri et al., 2023; Enticott et al., 2020

Source: our elaboration

5 Discussion and Conclusion

The shift to knowledge-based models in healthcare systems involves a shift away from fixed structures in favor of flexible systems. In this context, organizations are not simple units, but fluid structures shaped by the interactions between different actors. Research shows that an organization's strength and effectiveness depend not only on its resources, but also on the skills it develops over time. In particular, the ability to collaborate with others fosters teamwork, and digital skills facilitate the flow of information. Possessing the right organizational skills helps make strategic decisions in line with external regulations. Knowledge is created through collaboration between people. It is influenced by the organization's functioning and the coordination among its members. A key finding of the study is therefore that an organization's ability to adapt is linked to its self-knowledge. Organizations that understand their own strengths, weaknesses, and vulnerabilities are better at anticipating risks. Through knowledge management, past experiences can be transformed into learning, which reduces uncertainty in the decision-making process. The ability to change and adapt is crucial in this context because it allows organizations to better manage change. Furthermore, the study of relationships between organizations demonstrates that these can be both beneficial and problematic. Collaborative networks can contribute ideas and resources, but they can also create information gaps and bureaucratic problems. It is therefore essential to have systems that balance these different needs. From this perspective, this suggests that change should be allowed to occur spontaneously. This means investing not in individual skills, but in creating organizational systems that encourage continuous learning and teamwork.

6 Ethics Declaration

The authors declare that ethics clearance is not required for the research discussed in this document.

7 AI Declaration

The authors declare that artificial intelligence (AI) tools were used exclusively for linguistic translation, grammar correction, and formatting revision of the manuscript. No AI tools were used to generate text, develop arguments, or produce any substantive content included in this article. Furthermore, AI was not used for data analysis, interpretation of results, development of findings, or drawing conclusions. All intellectual contributions, theoretical development, methodological design, data analysis, and interpretation of results are solely attributable to the authors. The authors take full responsibility for the content of the manuscript.

References

Agranoff, R., & McGuire, M. (2003). Collaborative public management: New strategies for local governments. Georgetown University Press.

Al Amiri, N., Abdul Rahim, R. E., & Ahmed, G. (2020). Leadership Styles and Organizational Knowledge Management Activities: A Systematic Review. Gadjah Mada International Journal of Business.

Al Amiri, N., Abdul Rahim, R. E., & Ahmed, G. (2023). The Organizational Resources and Knowledge Management Capability: A Systematic Review. Journal of Information & Knowledge Management.
Alemu, B. A. (2025). Leveraging Knowledge Management for Sustainable Innovation: Advancing Public Health Leadership Interventions. Health Economics and Management Review.
Alexy, O., George, G., & Salter, A. J. (2013). Cui bono? The selective revealing of knowledge and its implications for innovative activity. Academy of management review, 38(2), 270-291.
Bagchi, M. (2021). Towards knowledge organization ecosystem (KOE). Cataloging & Classification Quarterly, 59(8), 740-756.
Barley, W. C. (2015). Anticipatory work: How the need to represent knowledge across boundaries shapes work practices within them. Organization Science, 26(6), 1612-1628.
Barry, D., Rerup, C., 2006. Going mobile: aesthetic design considerations from Calder and the constructivists. Organ. Sci. 17, 262–276.
Bate, P., Mendel, P., & Robert, G. (2008). Organizing for quality: The improvement journeys of leading hospitals in Europe and the United States. Radcliffe Publishing.
Clarysse, B., Wright, M., Bruneel, J., & Mahajan, A. (2014). Creating value in ecosystems: Crossing the chasm between knowledge and business ecosystems. Research Policy, 43(7), 1164–1176. https://doi.org/10.1016/j.respol.2014.04.014
Clarysse, B., Wright, M., Bruneel, J., & Mahajan, A. (2014). Creating value in ecosystems: Crossing the chasm between knowledge and business ecosystems. Research policy, 43(7), 1164- 1176.
Cobben, D., Ooms, W., Roijakkers, N., & Radziwon, A. (2022). Ecosystem types: A systematic review on boundaries and goals. Journal of Business Research, 142, 138-164.
Cormican, K., Meng, C., Sampaio, S., & Wu, Q. (2021). Towards Sustainable Knowledge Sharing Practices: An Analysis of Organizational Level Enablers. Sustainability.
Denis, J.-L., Langley, A., & Sergi, V. (2012). Leadership in pluralistic organizations. Academy of Management Annals, 6(1), 211–283. https://doi.org/10.1080/19416520.2012.667612
DI NAUTA, P., ; Sarno, D. (2017). Pratiche di co-creazione per il benessere dell'ecosistema di servizio sanitario. PROSPETTIVE IN ORGANIZZAZIONE, 7.
Dobusch, L., & Schoeneborn, D. (2015). Fluidity, identity, and organizationality: The communicative constitution of Anonymous. Journal of management studies, 52(8), 1005-1035.
Doherty, J., Johnson, M., & McPheron, H. (2021). Advancing health equity through organizational change: Perspectives from health care leaders. Health Care Management Review.
Enticott, J., Braaf, S., Johnson, A., Jones, A., & Teede, H. (2020). Leaders' perspectives on learning health systems: a qualitative study. BMC Health Services Research.
Ferlie, E., & Ongaro, E. (2015). Strategic management in public services organizations: Concepts, schools and contemporary issues. Routledge.
Ferlie, E., Crilly, T., Jashapara, A., & Peckham, A. (2012). Knowledge mobilization in healthcare: A critical review of health sector and generic management literature. Social Science & Medicine, 74(8), 1297–1304. https://doi.org/10.1016/j.socscimed.2011.11.042
Firmansyah, B. (2024). The Role of Systemic Enablers in Integrating Knowledge Management for Innovation. Data: Journal of Information Systems and Management.
Franzoni, C., & Sauermann, H. (2014). Crowd science: The organization of scientific research in open collaborative projects. Research Policy, 43(1), 1–20.
Frow, P. McColl-Kennedy J.R., Payne A. (2016). Co-creation practices: Their role in shaping a health care ecosystem. Industrial Marketing Management 56, 24-39.
Garud, R., Jain, S., & Tuertscher, P. (2008). Incomplete by design and designing for incompleteness. Organization studies, 29(3), 351-371.
Ghazinoory, S., Phillips, F., Afshari-Mofrad, M., & Bigdelou, N. (2021). Innovation lives in ecotones, not ecosystems. Journal of Business Research, 135, 572-580.
Helfat, C. E., Finkelstein, S., Mitchell, W., Peteraf, M. A., Singh, H., Teece, D. J., & Winter, S. G. (2007). Dynamic capabilities: Understanding strategic change in organizations. Blackwell Publishing.
Jacobides, M. G., Cennamo, C., & Gawer, A. (2018). Towards a theory of ecosystems. Strategic Management Journal, 39(8), 2255–2276. https://doi.org/10.1002/smj.2904
Järvi, K., Almpanopoulou, A., & Ritala, P. (2018). Organization of knowledge ecosystems: Prefigurative and partial forms. Research policy, 47(8), 1523-1537.
K. K., & Löfsten, H. (2023). Developing a knowledge ecosystem for large-scale research infrastructure. The Journal of Technology Transfer, 48(1), 441-467.
Kavalić, M., Nikolić, M., Radosav, D., Stanisavljev, S., & Pečujlija, M. (2021). Influencing Factors on Knowledge Management for Organizational Sustainability. Sustainability.
Kellogg, K.C., Orlikowski, W.J., Yates, J., 2006. Life in the trading zone: structuring coordination across boundaries in post-bureaucratic organizations. Organ. Sci. 17, 22–44
Kernick, D. (2004). Models, metaphors and paradigms: making sense of the world and the. Complexity and Healthcare Organization: A View from the Street, 3.
Kjellberg, H., Helgesson, C.F. (2007). On the nature of markets and their practices. Marketing Theory, 7(2), 137- 162. Laukka, E., Pölkki, T., & Kanste, O. (2022). Leadership in the context of digital health services: A concept analysis.

Journal of Nursing Management.
Leal, M., Leal, C., & Silva, R. (2025). Entrepreneurial Ecosystems and Knowledge Management: Systematic Literature Review and Bibliometric Analysis. Journal of the Knowledge Economy.
Mardian, M. T., Musadieq, M., Iqbal, M., & Ruhana, I. (2025). Understanding the Key Factors Influencing Knowledge Management in Business Environments: A Systematic Review. F1000Research.
Mondal, S., Rego, K., Talwar Kapoor, G., Wodnik, B. K., Law, M. P., & Di Ruggiero, E. (2025). Organizational Leadership Competencies for Public Health System Governance: A Scoping Review. Journal of Public Health Management and Practice.
Nicolini, D. (2012). Practice theory, work, and organization: An introduction. Oxford University Press.
Nonaka, I., & Takeuchi, H. (1995). The knowledge-creating company: How Japanese companies create the dynamics of innovation. Oxford University Press.
Öberg, C., & Lundberg, H. (2022). Mechanisms of knowledge development in a knowledge ecosystem. Journal of Knowledge Management, 26(11), 293-307.
Pellegrini, M., Ciampi, F., Marzi, G., & Orlando, B. (2020). The relationship between knowledge management and leadership: Mapping the field and providing future research avenues. Journal of Knowledge Management.
Perkmann, M., & Schildt, H. (2015). Open data partnerships between firms and universities: The role of boundary organizations. Research Policy, 44(5), 1133–1143.
Rabelo Neto, J., Figueiredo, C., Gabriel, B., & Valente, R. (2024). Factors for innovation ecosystem frameworks: Comprehensive organizational aspects for evolution. Technological Forecasting and Social Change.
Rosenkopf, L., Metiu, A., George, V.P., 2001. From the bottom up? Technical committee activity and alliance formation.
Admin. Sci. Q. 46, 748–772.
Royston, G., & Dick, P. (1998). Healthcare ecology. British Journal of Health Care Management, 4, 238-241.
Secundo, G., Toma, A., Schiuma, G., & Passiante, G. (2019). Knowledge transfer in open innovation: A classification framework for healthcare ecosystems. Business Process Management Journal, 25(1), 144-163.
Spanos, S., Leask, E., Patel, R., Datyner, M., Loh, E., & Braithwaite, J. (2024). Healthcare leaders navigating complexity: A scoping review of key trends in future roles and competencies. BMC Medical Education.
Teece, D. J. (2007). Explicating dynamic capabilities: The nature and microfoundations of (sustainable) enterprise performance. Strategic Management Journal, 28(13), 1319–1350. https://doi.org/10.1002/smj.640
Tessema, D. B. (2025). Pillars of organizational transformation and sustainability: Leadership, learning, culture, and knowledge management as key success factors in environmental public health. Health Economics and Management Review, 6(1), 1-21.
Tsoukas, H. (2009). A dialogical approach to the creation of new knowledge in organizations. Organization Science, 20(6), 941–957. https://doi.org/10.1287/orsc.1090.0435
Valkokari, K. (2015). (2015), Business, innovation, and knowledge ecosystems: How they differ and how to survive and thrive within them. Technology and Innovation Management Review, 5(8), 17–24.
Valkokari, K. (2015). Business, innovation, and knowledge ecosystems: How they differ and how to survive and thrive within them. Technology innovation management review, 5(8).
Van der Borgh, M., Cloodt, M., & Romme, A. G. L. (2012). Value creation by knowledge-based ecosystems: evidence from a field study. R&D Management, 42(2), 150-169.
Vesperi, W., Melina, A. M., Ventura, M., Coppolino, R., & Reina, R. (2021). Organizing knowledge transfer between university and agribusiness firms. Systems Research and Behavioral Science, 38(3), 321-329.
Wahid, M., Rahman, S., Yusuf, M., & Busyairi, M. H. (2024). Unlocking Institutional Capacity for Sustainable E- Government Transformation. Sinergi International Journal of Communication Sciences.
Wareham, J., Fox, P. B., & Cano Giner, J. L. (2014). Technology ecosystem governance. Organization Science, 25(4), 1195–1215

Designing Hybrid Human–AI Knowledge Ecosystems for Sustainable Value Creation

Stefania D'Aprile, Giovanni Spatola, Maria Zifaro, Giuseppe Ambrosio, Marianna Mancino
Universitas Mercatorum, Rome, Italy

Abstract

In the current context, characterized by increasing digitalization and a growing emphasis on sustainability, Artificial Intelligence (AI) is increasingly recognized as a key enabler of organizational performance and innovation. However, many organizations struggle to translate investments in AI into sustainable value due to knowledge fragmentation, weak governance mechanisms and limited integration between human capabilities and digital systems (Alavi & Leidner, 2001; Markus, 2001).

Building on the Knowledge Management (KM) and hybrid intelligence literature, this study proposes an operational framework for the design and governance of hybrid human–AI knowledge ecosystems aligned with ESG objectives. The framework is structured along the knowledge journey, encompassing the phases of planning, acquisition, and distribution and impact, and is empirically illustrated through a qualitative case study of Aulab Srl, with particular reference to the implementation of the Confluence platform as a Knowledge Management (KM) system.

The findings highlight the critical role of structured knowledge management practices as an enabling condition for effective AI adoption and sustainable value creation.

Keywords: hybrid intelligence; knowledge management; artificial intelligence; ESG; sustainable value

Paper type: Practical Paper

1 Introduction

The increasing prominence of sustainability concerns, coupled with the pervasive diffusion of digital technologies, is profoundly reshaping organizational value creation models. Within this evolving landscape, Artificial Intelligence (AI) has emerged as a strategic driver of innovation and performance enhancement. Despite its potential, the effective implementation of AI remains a significant challenge, as many organizations struggle to translate technological investments into sustainable outcomes.

As emphasized in the extant literature, these challenges are not primarily attributable to technological limitations, but rather to organizational and cognitive issues, including knowledge fragmentation, weak governance mechanisms, and insufficient integration between human capabilities and digital systems (Alavi & Leidner, 2001; Markus, 2001). In this regard, Knowledge Management (KM) constitutes a critical lever for addressing such complexities, as it enables the systematic structuring, sharing, and reuse of organizational knowledge.

Furthermore, the advent of digital technologies and advanced algorithms is fundamentally transforming knowledge management practices as well as decision-making processes (Faraj et al., 2018). From this perspective, AI can be conceptualized as a socio-technical phenomenon, whose value is contingent upon its effective integration within organizational systems and processes (Jarrahi, 2018). In parallel, the paradigm of hybrid intelligence underscores that superior outcomes are achieved through the complementary design of human and artificial capabilities, rather than through substitution logics (Dellermann et al., 2019).

Notwithstanding these advancements, the literature still provides limited operational guidance on how to design and govern hybrid human–AI knowledge ecosystems oriented toward sustainability and aligned with ESG objectives (Mikalef et al., 2023). Against this backdrop, the present study aims to develop an operational framework that integrates Knowledge Management, hybrid intelligence, and ESG principles along the knowledge journey.

2 Theoretical Background

2.1 Knowledge Management and Value Creation

The Knowledge Management (KM) literature widely recognizes knowledge as a central strategic resource for achieving competitive advantage and fostering organizational innovation. Seminal contributions emphasize that the ability to generate, share and reuse knowledge constitutes a fundamental determinant of organizational performance (Alavi & Leidner, 2001; Markus, 2001).

From this perspective, KM can be conceptualized as an integrated set of practices, processes, and systems aimed at facilitating the integration and leveraging of distributed knowledge within the organization. Theoretically, KM is grounded in the knowledge-based view of the firm, which conceptualizes organizations as systems for the integration of specialized knowledge (Grant, 1996). Within this framework, value creation depends on the organization's capacity to coordinate such knowledge and transform it into distinctive organizational capabilities.

More recently, the literature has increasingly emphasized the digital dimension of KM, highlighting how collaborative platforms and advanced information systems contribute to making knowledge more accessible, scalable, and reusable (Majchrzak et al., 2016). However, the adoption of these technologies necessitates the implementation of appropriate governance mechanisms to mitigate risks such as information overload and knowledge fragmentation.

2.2 Artificial Intelligence and the Transformation of Knowledge Processes

The advent of Artificial Intelligence has introduced a profound shift in knowledge management processes, redefining the ways in which organizations generate, analyze and utilize information. As highlighted by Faraj et al. (2018), machine learning algorithms do not merely automate existing tasks; rather, they fundamentally transform organizational practices, reshaping the very nature of work and decision-making processes. In this context, AI is increasingly conceptualized as a socio-technical phenomenon, whose value emerges from the interaction between technological systems, human actors, and organizational structures (Jarrahi, 2018). Recent studies further emphasize that the adoption of AI necessitates new forms of organizational design and coordination, capable of integrating advanced analytical capabilities with complex decision-making processes (Raisch & Krakowski, 2021). At the same time, the literature underscores that the effectiveness of AI is highly contingent upon the quality of the data and knowledge on which it is grounded. In the absence of a robust Knowledge Management infrastructure, AI systems risk generating biased or unreliable outputs, thereby amplifying existing inefficiencies and distortions (Davenport & Ronanki, 2018).

2.3 Hybrid Intelligence: Complementarity Between Humans and AI

The concept of hybrid intelligence has emerged in response to the growing recognition of the inherent limitations of both human and artificial intelligence. From this perspective, value creation stems from the design of systems in which humans and AI operate in a complementary manner, combining computational capabilities with interpretative and cognitive skills (Dellermann et al., 2019).

The literature highlights that AI excels in processing large volumes of data, pattern recognition, and scalability, whereas human actors contribute critical thinking, ethical judgment and contextual understanding (Jarrahi, 2018). This complementarity is particularly salient in decision-making contexts characterized by uncertainty and ambiguity, where purely automated approaches prove insufficient.

Recent contributions further emphasize the importance of organizational design in shaping hybrid intelligence configurations, showing that value emerges from the effective orchestration of human–AI interactions within organizational processes and routines (Raisch & Krakowski, 2021). This entails a shift from substitution-based logics toward augmentation-oriented approaches, in which technology enhances rather than replaces human capabilities.

2.4 ESG, Sustainability and Knowledge Governance

The increasing focus on ESG (Environmental, Social, and Governance) objectives has introduced new challenges for knowledge and data management. In particular, achieving these objectives requires high levels of transparency, reliability and traceability of information, as well as heightened attention to ethical considerations and accountability (Floridi et al., 2018).

AI can play an enabling role in supporting sustainability goals, for instance through the analysis of environmental data or the monitoring of social performance. However, the literature also highlights the risks

associated with AI adoption, including algorithmic bias, opacity, and potential misalignment with organizational values (Vinuesa et al., 2020). These challenges underscore the need for robust governance mechanisms capable of ensuring the responsible and transparent use of such technologies.

In this context, Knowledge Management assumes a central role as an enabling infrastructure for sustainability, as it supports the structuring and accessibility of the knowledge required for informed and responsible decision-making. As recent studies suggest, knowledge management is deeply intertwined with digital transformation processes and represents a key pillar for the development of sustainable and resilient organizations (Heisig et al., 2022).

2.5 Research Gap and Study Contribution

Despite significant advancements across the individual streams of research, the literature remains fragmented, with limited integration among Knowledge Management, Artificial Intelligence and sustainability. In particular, there is a lack of operational models that provide clear guidance on how to design and govern hybrid human–AI knowledge ecosystems capable of supporting sustainable value creation (Mikalef et al., 2023).

This study positions itself within this research gap by proposing an integrated framework that connects Knowledge Management, hybrid intelligence and ESG principles along the knowledge journey, thereby contributing to the development of a unified theoretical and practical perspective. In doing so, it responds to the need to move beyond fragmented approaches and offers actionable insights for the design of knowledge ecosystems capable of sustaining innovation and long-term sustainability.

3 Methodology

This study adopts a qualitative approach based on a case study research design, which is particularly suitable for investigating complex, emerging and highly contextualized phenomena such as the design and governance of hybrid human–AI knowledge ecosystems. Consistent with methodological literature, the case study approach enables an in-depth understanding of organizational dynamics and socio-technical processes, proving especially appropriate when the boundaries between the phenomenon and its context are not clearly delineated (Yin, 2018).

Furthermore, in line with Eisenhardt's (1989) approach, the case study facilitates the generation of theoretical insights grounded in empirical analysis, thereby contributing to the development of conceptually robust and practice-based models. Case selection follows a theoretical sampling logic, identifying Aulab Srl as a particularly informative and revelatory context.

As a knowledge-intensive EdTech organization, Aulab represents a relevant empirical setting for examining the role of Knowledge Management as an enabling factor for the adoption of Artificial Intelligence and for supporting decision-making processes aligned with sustainability and ESG objectives. Moreover, the case provides a privileged opportunity to observe the interaction between knowledge management practices, digital infrastructures and governance mechanisms within a highly dynamic and innovation-driven environment.

Data collection was conducted through a combination of qualitative sources to ensure triangulation and analytical robustness. Specifically, the analysis draws on internal documentation, content available within the Confluence platform, organizational materials, and observations related to knowledge structuring and utilization processes. This plurality of data sources enabled a comprehensive reconstruction of Knowledge Management practices and governance mechanisms, as well as an analysis of their implications for the development of hybrid intelligence configurations.

The use of documentary and digital data is particularly consistent with the study's objectives, as it allows for direct observation of how knowledge is codified, organized and made accessible within the organizational ecosystem. The empirical analysis focuses on three main dimensions: (i) the design of the knowledge ecosystem, (ii) governance mechanisms and (iii) the role of digital platforms as enabling infrastructures for human–AI integration.

Through an iterative process between empirical evidence and the relevant literature, the study develops an interpretive framework that integrates Knowledge Management, hybrid intelligence and ESG principles along the knowledge journey.

Notwithstanding its contributions, the study presents several limitations. First, the qualitative nature of the research and the reliance on a single case study limit the generalizability of the findings, which should be interpreted in analytical rather than statistical terms. Second, the focus on a specific organization operating in the EdTech sector may not fully capture dynamics present in other organizational contexts. Additionally, the predominance of documentary data may not entirely capture implicit or informal dimensions of knowledge management, such as tacit practices and relational dynamics.

Future research could address these limitations by extending the analysis through comparative multi-case studies or by integrating qualitative and quantitative methodologies, thereby enhancing the validity and transferability of the findings.

4 A Framework for Hybrid Knowledge Ecosystems

The proposed framework is situated at the intersection of Knowledge Management literature, recent advancements in hybrid intelligence and the growing body of research on sustainability and ESG objectives, offering an integrated perspective on these domains along the knowledge journey (Heisig et al., 2022). Within this configuration, knowledge is conceptualized not as a static asset, but as a dynamic and cumulative process that unfolds across multiple stages of generation, codification, sharing and application, in continuous interaction with organizational structures, digital technologies and governance mechanisms.

From a theoretical standpoint, Knowledge Management assumes the role of an enabling socio-technical infrastructure, within which practices, routines and organizational mechanisms are articulated to systematize knowledge. In this view, the adoption of shared standards, taxonomies and templates is not merely an operational choice, but a mechanism for reducing ambiguity and enhancing organizational coordination, consistent with theories of knowledge integration and knowledge reuse (Alavi & Leidner, 2001; Markus, 2001). At the same time, the definition of roles and responsibilities, such as knowledge owners, contributes to the establishment of distributed accountability, strengthening processes of validation, updating, and knowledge quality assurance. This perspective positions KM as an organizational architecture capable of supporting complex decision-making processes and enabling the scalability of knowledge practices.

Within this infrastructure, the framework incorporates the paradigm of hybrid intelligence, conceptualized as an emergent configuration of complementarity between human and artificial capabilities. In line with a socio-technical perspective, Artificial Intelligence is not viewed as a substitute for human agency, but rather as an augmentative element that extends cognitive capabilities, particularly in terms of processing large volumes of data, identifying patterns, and providing analytical support (Dellermann et al., 2019; Jarrahi, 2018). Conversely, human actors retain a central role in sensemaking processes, meaning construction, and the exercise of ethical judgment, which are especially critical in contexts characterized by high levels of uncertainty and ambiguity.

Accordingly, the framework advocates for the intentional design of human–AI interactions oriented toward augmentation rather than substitution, emphasizing that value creation emerges from the effective orchestration of these interactions within organizational processes.

A further distinctive feature of the framework lies in the explicit integration of ESG objectives, which introduces a normative and value-based dimension into knowledge management. In this respect, the quality, traceability, and reliability of knowledge are not merely operational requirements, but become essential conditions for ensuring transparency, accountability, and alignment between organizational decisions and socio-environmental impacts (Floridi et al., 2018; Vinuesa et al., 2020). The integration of ESG within the knowledge journey thus entails a redefinition of knowledge evaluation criteria, extending beyond efficiency and effectiveness to encompass ethical and sustainability considerations.

Overall, the proposed framework conceptualizes hybrid knowledge ecosystems as complex adaptive systems, in which technological, organizational, and institutional dimensions co-evolve. Within this perspective, governance assumes a central role in orchestrating these dimensions, ensuring alignment between strategy, knowledge practices, and sustainability objectives.

The theoretical contribution of the framework lies in its ability to integrate traditionally separate domains—Knowledge Management, Artificial Intelligence, and sustainability—offering a unified interpretive lens through which to understand how organizations can design and govern knowledge ecosystems oriented toward long-term sustainable value creation.

5 The Knowledge Journey

The proposed framework unfolds across three interrelated phases of the knowledge journey, conceptualized as a dynamic and continuous process through which knowledge is not only generated and managed, but also progressively integrated into decision-making processes and value creation systems. The three phases—planning, acquisition, structuring, distribution and impac, should not be interpreted as rigidly sequential stages, but rather as interdependent and recursive dimensions, characterized by continuous feedback loops and organizational learning processes.

The **planning phase** represents the stage in which knowledge is recognized and formalized as a strategic resource. During this phase, organizations identify critical knowledge domains in alignment with business objectives and ESG orientations, while simultaneously defining the informational requirements necessary to

support decision-making processes. Planning also entails the design of knowledge governance mechanisms, including the definition of roles, responsibilities and dedicated organizational structures. In this perspective, top management plays a pivotal role in ensuring alignment between strategy, Knowledge Management and AI initiatives, thereby guiding the creation and utilization of knowledge toward sustainable value objectives. This phase can thus be understood as a moment of *strategic sensemaking*, in which the cognitive and informational priorities of the organization are established.

The **acquisition and structuring phase** concerns the processes through which knowledge is collected, codified, and made accessible. At this stage, organizations face the challenge of transforming tacit knowledge, often distributed across individuals and teams, and explicit knowledge embedded in documents, procedures and systems into structured and shared forms. The use of digital Knowledge Management platforms is critical in this regard, as it enables the systematization of content through the adoption of shared standards, templates and taxonomies (Majchrzak et al., 2016). Importantly, this process is not purely technical; it also entails a social dimension, involving the active participation of organizational actors and the development of shared practices for documentation and continuous updating. The quality of this phase directly influences the organization's ability to develop a coherent, reliable, and reusable knowledge base, which constitutes a prerequisite for the effective integration of AI.

The **distribution and impact phase** represents the stage in which structured knowledge is activated within organizational processes, contributing to value creation. In this phase, knowledge is disseminated, reinterpreted and reused within both operational and strategic decision-making processes, often supported by Artificial Intelligence tools that enhance analytical and predictive capabilities. AI can facilitate access to knowledge, support data analysis and improve decision quality; however, its effectiveness remains contingent upon the robustness of the underlying knowledge infrastructure. At the same time, this phase encompasses mechanisms for monitoring and evaluating outcomes, both in terms of organizational performance and alignment with ESG objectives. This enables the closure of the knowledge journey cycle, generating feedback that feeds back into the planning phase and fosters continuous learning processes.

Overall, the analysis of the knowledge journey highlights that sustainable value creation depends on the organization's ability to orchestrate these three phases in an integrated manner, overcoming fragmented approaches to knowledge management. From this perspective, the knowledge journey should not be conceived merely as a sequence of activities, but as a dynamic infrastructure through which organizations continuously build, transform and mobilize knowledge in response to the challenges posed by digitalization and sustainability.

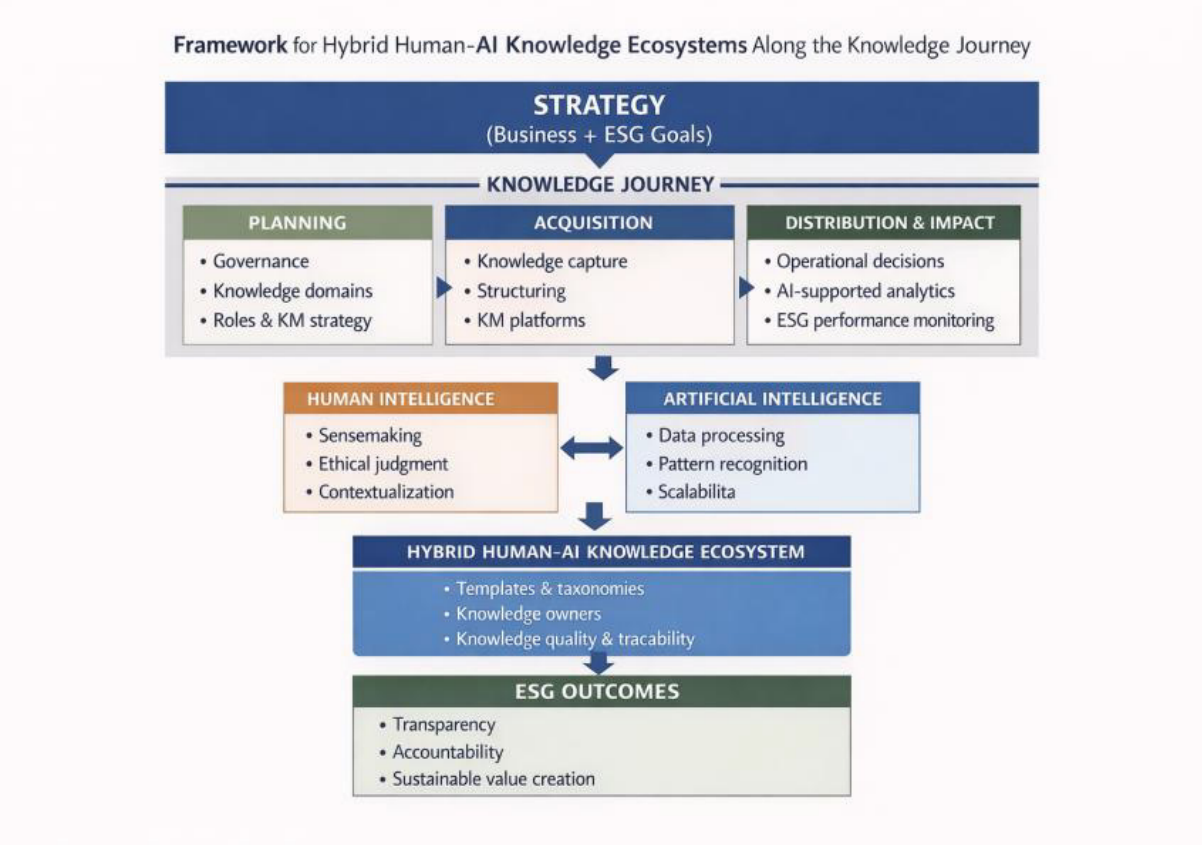


Figure 1: Framework for hybrid Human-AI knowledge ecosystem along the knowledge journey

6 Case Study: Aulab and the Implementation of Confluence

The case of Aulab Srl represents a particularly insightful context for examining how the systematic structuring of knowledge can act as a foundational condition for the evolution toward hybrid human–AI knowledge ecosystems. As a knowledge-intensive EdTech organization, Aulab operates in an environment where the ability to generate, update, share, and reuse knowledge constitutes a critical resource for both operational efficiency and the quality of decision-making processes.

Within this context, the implementation of the Confluence platform should not be interpreted merely as the adoption of a technological tool, but rather as an organizational intervention aimed at enhancing the visibility, accessibility, governability and integration of knowledge within business processes. More specifically, Confluence was introduced as a Knowledge Management infrastructure designed to collect and organize the organization's informational assets within thematic spaces aligned with its core knowledge domains, including educational processes, operational procedures, technological tools and elements related to organizational culture.

This structuring enabled the organization to move beyond a fragmented logic of information management, often dispersed across isolated documents, informal communications and individually held knowledge, toward a progressive codification of knowledge into shared and reusable forms. The adoption of standardized templates, common taxonomies and consistent documentation practices further reinforced this process, enhancing the comparability, updatability, and retrievability of content.

A particularly relevant aspect emerging from the case concerns the governance dimension of knowledge. The introduction of specific roles, such as knowledge owners, contributed to clarifying responsibilities related to the creation, validation, maintenance, and updating of content. In this sense, the platform not only improved information storage, but also fostered the development of a system of distributed accountability, in which knowledge quality becomes an explicit object of organizational oversight. This is especially critical in dynamic environments such as Aulab, where the rapid evolution of educational content, operational practices and digital tools requires continuous mechanisms of revision and alignment.

From an organizational outcomes perspective, the implementation of Confluence significantly reduced information fragmentation and the risk of loss of tacit knowledge, particularly that embedded in the day-to-day experience of individual actors. The systematization of content improved accessibility, quality, and traceability of information, facilitating not only internal coordination but also organizational learning and onboarding processes. In other words, the platform supported the transition from dispersed and localized knowledge toward more explicit, shared and potentially scalable knowledge.

From a theoretical standpoint, the Aulab case illustrates how the availability of a structured knowledge base constitutes a critical precondition for the effective adoption of Artificial Intelligence. AI systems derive value from the presence of organized, reliable, and semantically coherent data and content; in the absence of such an infrastructure, there is a risk of feeding automated processes with incomplete, inconsistent or weakly validated knowledge. In this sense, Confluence can be interpreted as an enabling infrastructure for future hybrid intelligence configurations, in which AI tools support the analysis, retrieval, and valorization of knowledge, while human actors continue to play a decisive role in interpretation, contextualization, and judgment.

Finally, the case also highlights a clear connection with ESG objectives. Increased transparency of information flows, traceability of sources and modifications, and the ability to reconstruct responsibilities and decision-making processes strengthen the conditions for more responsible and sustainable governance. From this perspective, the platform contributes not only to organizational efficiency, but also to the development of a knowledge ecosystem capable of supporting more informed, verifiable and ESG-aligned decision-making processes oriented toward long-term value creation.

7 Discussion

The findings of this study confirm that the effective adoption of Artificial Intelligence cannot be interpreted as a purely technological process, but must be framed within a broader transformation of knowledge management practices. In particular, the value of AI is shown to be highly contingent upon the quality, coherence, and governance of the knowledge on which it is grounded. In the absence of structured Knowledge Management practices, the introduction of AI risks amplifying existing inefficiencies, reproducing informational fragmentation and organizational inconsistencies, in line with prior research on the digital transformation of knowledge processes (Faraj et al., 2018).

The case analysis further highlights the centrality of governance in the configuration of effective knowledge ecosystems. The introduction of formalized roles, shared standards and integrated platforms contributes to reducing ambiguity and strengthening organizational coordination, thereby enabling the development of

knowledge that is more reliable, accessible and reusable. In this sense, governance should be understood not as a static set of rules, but as a dynamic orchestration mechanism that aligns actors, technologies, and practices.

From a hybrid intelligence perspective, the findings suggest that value creation does not stem from the mere adoption of AI tools, but from the intentional design of complementary configurations between human and artificial capabilities. While AI extends the analytical capacity of organizations, human actors retain a central role in processes of interpretation, contextualization and ethical evaluation. Consequently, the effectiveness of hybrid intelligence depends on the organization's ability to embed these complementary dimensions within established practices, rather than relying solely on technological availability (Dellermann et al., 2019; Jarrahi, 2018).

An additional contribution of this study concerns the integration of ESG objectives, which emerges as a structural rather than peripheral dimension. The availability of traceable and reliable knowledge enhances transparency and accountability, which are foundational elements of sustainable governance. In this regard, Knowledge Management plays a pivotal role not only in enabling AI adoption, but also in supporting the alignment between organizational performance and socio-environmental impacts (Floridi et al., 2018; Vinuesa et al., 2020).

Overall, the theoretical contribution of this work lies in its ability to integrate three traditionally distinct streams of literature, Knowledge Management, hybrid intelligence, and sustainability, into a unified perspective grounded in the knowledge journey. This integrative approach helps to overcome the fragmentation of existing research and provides a systemic lens for understanding the design and governance of hybrid knowledge ecosystems.

8 Managerial and Policy Implications

From a managerial perspective, the findings suggest that organizations should treat Knowledge Management as a strategic priority rather than a supporting activity. In particular, before embarking on advanced AI adoption initiatives, it is essential to invest in the structuring of knowledge through shared standards, taxonomies and integrated digital systems, in order to ensure the quality and reliability of data.

Moreover, there is a clear need to develop well-defined and formalized knowledge governance models, grounded in the specification of roles and responsibilities. The introduction of roles such as knowledge owners can strengthen accountability mechanisms and ensure the continuous updating of content, thereby reducing the risk of obsolescence and informational inconsistency.

A further critical aspect concerns the design of human–AI interactions. Organizations should move away from substitution-oriented approaches and instead adopt complementarity-based logics, designing decision-making processes in which the analytical capabilities of AI are effectively integrated with the interpretative and strategic competencies of human actors. This also entails the development of specific organizational capabilities, not only related to the use of technologies, but also to their meaningful integration within work processes.

Finally, the integration of ESG objectives requires a rethinking of knowledge management systems in terms of transparency, traceability and accountability. In this regard, Knowledge Management platforms can play a crucial role in supporting more informed decision-making practices and in ensuring alignment between corporate strategies and long-term socio-environmental impacts.

From a policy-making perspective, the findings highlight the need to foster institutional ecosystems that support the integration of Knowledge Management, Artificial Intelligence and sustainability. In particular, it's essential to promote the adoption of common standards for knowledge management and sharing, in order to enhance interoperability across organizations and facilitate the responsible use of data.

Furthermore, public policies should support the development of hybrid competencies, combining digital literacy, analytical capabilities and ethical awareness, to better prepare individuals and organizations to operate in contexts characterized by increasing human–AI integration. This implies investments in education, training, and reskilling programs focused on knowledge management and the responsible use of AI.

An additional area of intervention concerns the development of regulatory frameworks and guidelines aimed at ensuring transparency, accountability, and reliability in AI-based systems, particularly in relation to ESG objectives. Strengthening knowledge and data traceability practices can contribute to improving the quality of both public and private decision-making, while mitigating risks associated with algorithmic bias and process opacity.

Finally, policy makers can play a key role in promoting the diffusion of digital infrastructures and collaborative platforms that facilitate knowledge sharing between public and private actors, thereby fostering the development of innovation ecosystems oriented toward sustainability.

Overall, these interventions can contribute to establishing the institutional conditions necessary for the development of hybrid knowledge ecosystems capable of generating sustainable value in the long term.

9 Conclusions

This study demonstrates that the design of hybrid human–AI knowledge ecosystems is not merely an evolutionary option, but a fundamental strategic lever for organizations operating in contexts characterized by increasing complexity, digitalization and mounting sustainability pressures. By integrating Knowledge Management, hybrid intelligence and ESG objectives along the knowledge journey, the proposed framework highlights how sustainable value creation increasingly depends on the systemic orchestration of knowledge, technology and governance.

From this perspective, the contribution of the study extends beyond a purely technological understanding of AI, emphasizing that true competitive advantage lies in the ability to structure, make accessible and activate organizational knowledge within hybrid decision-making processes. Organizations that invest in robust Knowledge Management infrastructures and governance models are better positioned not only to adopt AI effectively, but also to direct its use toward transparency, accountability, and long-term value creation.

From a managerial standpoint, the findings point to a necessary paradigm shift: AI adoption should not be driven solely by technological considerations, but embedded within a broader knowledge management strategy. This entails rethinking roles, processes and organizational practices to foster the integration of human and artificial capabilities, thereby transforming knowledge into a key driver of sustainable innovation.

From an academic perspective, the study contributes to advancing the debate on hybrid intelligence by moving beyond a predominantly conceptual focus toward a more practice-oriented approach. In doing so, it offers an integrated framework that bridges theory and practice, addressing a gap in the existing literature while opening new avenues for research on the interplay between knowledge ecosystems, emerging technologies and sustainability.

Ultimately, the future of value creation will not depend solely on the adoption of Artificial Intelligence, but on organizations' ability to design knowledge ecosystems in which human and AI capabilities co-evolve, positioning knowledge as a strategic infrastructure for truly scalable and enduring sustainability. In this sense, the proposed framework not only fills a gap in existing literature, but also opens up new research avenues on the interaction between knowledge ecosystems, emerging technologies, and sustainability. In particular, future research could further investigate the empirical applicability of the model across different organizational contexts through comparative multi-case studies, as well as integrate quantitative methods to assess the impact of hybrid knowledge ecosystems on ESG performance. Additional studies could also explore the role of emerging technologies, such as generative AI, in enhancing knowledge creation and decision-making processes, as well as examine the ethical and governance implications associated with these developments.

10 Ethics Declaration

The authors declare that this study was conducted in accordance with the ethical standards generally accepted for scientific research. Where applicable, all experimental procedures were approved by the appropriate ethics committees and comply with national and international regulations. The authors declare no conflicts of interest.

11 AI Declaration

During the preparation of this manuscript, generative Artificial Intelligence tools were used to assist with writing and editing. Specifically, the language model ChatGPT developed by OpenAI (GPT-4 version) was employed to improve textual clarity, harmonize linguistic style, and provide suggestions for rephrasing complex sections, under the direct supervision and final approval of the human authors. No part of the content was generated autonomously by AI without human oversight. The authors take full responsibility for the scientific content and conclusions presented.

References

Alavi, M., & Leidner, D. E. (2001). Review: Knowledge management and knowledge management systems: Conceptual foundations and research issues. *MIS Quarterly*, 25(1), 107–136.

Brynjolfsson, E., & McAfee, A. (2017). *Machine, platform, crowd: Harnessing our digital future*. W. W. Norton & Company.

Davenport, T. H., & Ronanki, R. (2018). Artificial intelligence for the real world. *Harvard Business Review*, 96(1), 108–116.

Dellermann, D., Ebel, P., Söllner, M., & Leimeister, J. M. (2019). Hybrid intelligence. *Business & Information Systems Engineering*, 61(5), 637–643.

Dignum, V. (2019). *Responsible artificial intelligence: How to develop and use AI in a responsible way*. Springer.

Eisenhardt, K. M. (1989). Building theories from case study research. *Academy of Management Review*, 14(4), 532–550.

European Commission. (2019). *Ethics guidelines for trustworthy AI*. Publications Office of the European Union.

Faraj, S., Pachidi, S., & Sayegh, K. (2018). Working and organizing in the age of the learning algorithm. *Information and Organization*, 28(1), 62–70.

Floridi, L., Cowls, J., Beltrametti, M., et al. (2018). AI4People: An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines*, 28(4), 689–707.

Grant, R. M. (1996). Toward a knowledge-based theory of the firm. *Strategic Management Journal*, 17(S2), 109–122.

Heisig, P., Schulte, T., & Vanhaverbeke, W. (2022). Digital transformation and knowledge management. *Journal of Knowledge Management*, 26(3), 454–476.

Jarrahi, M. H. (2018). Artificial intelligence and the future of work: Human–AI symbiosis in organizational decision making. *Business Horizons*, 61(4), 577–586.

Kane, G. C., Palmer, D., Phillips, A. N., Kiron, D., & Buckley, N. (2015). Strategy, not technology, drives digital transformation. *MIT Sloan Management Review*.

Majchrzak, A., Markus, M. L., & Wareham, J. (2016). Designing for digital transformation: Lessons for information systems research from the study of ICT and societal challenges. *MIS Quarterly*, 40(2), 267–277.

Markus, M. L. (2001). Toward a theory of knowledge reuse: Types of knowledge reuse situations and factors in reuse success. *Journal of Management Information Systems*, 18(1), 57–93.

Mikalef, P., Conboy, K., Lundström, J., & Popovič, A. (2023). Thinking responsibly about responsible AI. *Information Systems Frontiers*, 25, 1–17.

Nonaka, I., & Takeuchi, H. (1995). *The knowledge-creating company: How Japanese companies create the dynamics of innovation*. Oxford University Press.

OECD. (2019). *OECD principles on artificial intelligence*. OECD Publishing.

Pan, S. L., Pan, G., Leidner, D. E., et al. (2022). Human–AI collaboration: A review and research agenda. *Information Systems Journal*, 32(4), 699–733.

Raisch, S., & Krakowski, S. (2021). Artificial intelligence and management: The automation–augmentation paradox. *Academy of Management Review*, 46(1), 192–210.

Samsul Alam, A., Khaled Aminul Islam, B., Repon Miah, C., Sumaiya Afrin, D., Ariful Islam, B., Md. Rakibul Hoque, M., & Mahmud, I. (2025). Key success factors of knowledge management systems for optimizing organizational performance: A PRISMA-based systematic literature review. *Business and Management*, 11(16), e44097.

Spagnoletti, P., Resca, A., & Lee, G. (2021). A design theory for digital platforms supporting online communities. *Journal of Knowledge Management*, 25(6), 1439–1463.

Vinuesa, R., et al. (2020). The role of artificial intelligence in achieving the Sustainable Development Goals. *Nature Communications*, 11(1), 233.

Wilson, H. J., & Daugherty, P. R. (2018). *Human + machine: Reimagining work in the age of AI*. Harvard Business Review Press.

Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). Sage Publications.

Transforming Cyberbullying Prevention Guidelines into Actionable Workflows through AI-based Detection and BPMN Modeling

Alessandro Massaro, Giuseppe Loseto, Filippo Gramegna, Giovanni Schiuma, Francesco Santarsiero
LUM "Giuseppe Degennaro" University, Casamassima, Italy

Abstract

Cyberbullying represents a pervasive and evolving threat within digital environments, exacerbated by the scale, anonymity, and immediacy of online interactions. While recent advances in Artificial Intelligence (AI), particularly in Natural Language Processing (NLP), have significantly improved the ability to automatically detect harmful content, a critical gap persists between detection and effective intervention. Existing approaches primarily focus on identifying cyberbullying instances, but they often lack mechanisms to operationalize prevention guidelines and translate them into consistent, actionable responses. This paper addresses this gap by proposing a novel framework that integrates AI-based cyberbullying detection with Business Process Model and Notation (BPMN) to transform textual prevention guidelines into executable workflows. The framework leverages AI models to analyze user-generated content and identify potential cyberbullying signals, which are then used to trigger BPMN-based processes representing structured intervention strategies. A key contribution of this work lies in the systematic mapping of unstructured policy documents into BPMN constructs, allowing organizations to bridge the gap between high-level guidelines and operational procedures. The proposed approach is illustrated through a detailed modeling of cyberbullying prevention guidelines, demonstrating how complex policies can be translated into actionable pipelines. By combining the strengths of AI-driven detection and process-oriented modeling, this work contributes to the development of more effective and scalable cyberbullying prevention systems. Ultimately, the framework aims to support the creation of safer digital environments by enabling not only the identification of harmful behavior but also the consistent execution of appropriate and timely responses, paving the way for future research at the intersection of AI, process management, and digital safety.

Keywords: BPMN, Cyberbullying, Artificial Intelligence, Process Modeling.

Paper type: Academic Research Paper

1 Introduction and Motivation

Cyberbullying has emerged as a pervasive cyber-social security challenge, producing significant psychological, social, and economic impacts across digital platforms (Giumetti and Kowalski, 2022). In particular, the rapid proliferation of digital communication platforms has fundamentally reshaped social interaction, enabling unprecedented connectivity while simultaneously amplifying harmful behaviors. Unlike traditional forms of bullying, cyberbullying is pervasive, persistent, and often difficult to detect due to the scale and velocity of online content generation. Social media platforms, messaging services, and online communities generate vast streams of user-generated data, making manual moderation approaches increasingly insufficient. As a result, the need for automated, scalable, and context-aware solutions to identify and mitigate cyberbullying has become a pressing concern for researchers, policymakers, and platform providers alike. Over the past decade, advances in Artificial Intelligence (AI), particularly in Natural Language Processing (NLP) and Machine Learning (ML), have enabled the development of sophisticated cyberbullying detection models. These systems can analyze textual, contextual, and behavioral signals to identify potentially harmful interactions with increasing accuracy. However, despite significant progress in detection capabilities, a critical gap remains between identifying harmful content and operationalizing appropriate intervention strategies. In practice, organizations often rely on textual guidelines, policy documents, or best practices for cyberbullying prevention, which are inherently abstract, ambiguous, and difficult to implement consistently across complex socio-technical systems (Rahman et al., 2025). This disconnect highlights a key limitation in current approaches: the lack of structured mechanisms to translate high-level prevention guidelines into executable processes.

While AI models can flag potentially harmful content, they do not inherently provide a standardized framework for deciding what actions should follow, how responsibilities should be distributed among stakeholders, or how interventions should adapt to different contexts. Consequently, responses to cyberbullying incidents may be inconsistent, delayed, or ineffective, undermining the overall goal of safeguarding users. To address this challenge, there is a growing need for methodologies that bridge the gap between detection and action by formalizing intervention workflows. Business Process Model and Notation (BPMN) offers a well-

established standard for modeling, analyzing, and automating organizational processes (Rosa et al., 2026). By representing activities, decision points, and information flows in a structured and visual manner, BPMN enables the transformation of informal guidelines into precise, executable workflows. Despite its widespread adoption in business process management, its potential in the domain of cyberbullying prevention remains largely unexplored.

This paper proposes a novel framework that integrates AI-based cyberbullying detection with BPMN modeling to transform textual prevention guidelines into actionable workflows. The core idea is to create a seamless pipeline in which detection outputs from AI models serve as triggers for BPMN-defined processes, enabling systematic and context-aware responses to cyberbullying incidents. By formalizing intervention strategies as process models, the framework aims to ensure consistency, transparency, and adaptability in handling harmful online behavior. The motivation behind this work is twofold. First, it seeks to enhance the practical applicability of AI-based detection systems by embedding them within structured decision-making processes. Second, it aims to support organizations, such as social media platforms, educational institutions, and regulatory bodies, in operationalizing their cyberbullying prevention policies in a clear and scalable manner. Ultimately, this approach contributes to the development of more resilient digital ecosystems, where harmful behaviors can be not only detected but also addressed through well-defined and effective intervention workflows.

The remainder of this paper is structured as follows. Section 2 presents the proposed framework architecture, detailing the integration between AI-based cyberbullying detection components and BPMN-driven process orchestration. Section 3 focuses on the modeling of cyberbullying prevention guidelines, illustrating how textual policies can be systematically translated into structured and executable BPMN workflows. Finally, Section 4 concludes the paper by summarizing the main contributions and outlining directions for future research.

2 Framework Architecture

The proposed architecture, depicted in Figure 1, integrates multiple components to support cyberbullying analysis. Specifically, it is organized into four main modules: data preparation; model building; platform integration; and end-user support.

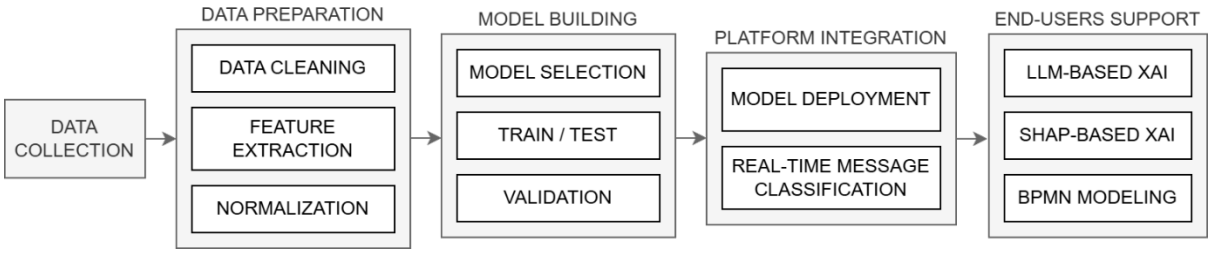


Figure 1 - Framework Architecture

The data preparation module is responsible for importing, cleaning, and preparing the reference dataset, to improve the quality and representativeness of the input data. The outcome of this phase is a structured and standardized dataset ready for model training and real-time inference. The core analytical component of the framework applies supervised learning algorithms for the classification of messages as cyberbullying or non-cyberbullying. Depending on the experimental configuration, both traditional ML models and Deep Learning approaches can be employed. The models are trained on pre-processed data and optimized using basic classification metrics such as accuracy, precision and recall. This module is designed for scalability and real-time execution to support continuous message analysis. The framework also integrates a web interface acting as the user-facing component of the system, simulating a social networking environment in which users can post, view, and analyze messages. It can integrate the reference classification model to provide immediate feedback on message content, displaying whether a given message is identified as cyberbullying. Finally, the eXplainable AI (XAI) module (see Loseto et al., 2026) enhances the transparency and interpretability of the classification process. It combines two complementary approaches: a local Large Language Model (LLM) capable of generating natural language explanations to communicate model reasoning to non-technical users, and the SHAP library, which provides quantitative and visual insights into the contribution of individual features.

XAI represents an additional element for increasing trust and transparency in automated decision-making systems. In the context of cyberbullying detection and broader cybersecurity applications, the ability to understand and communicate why a system produced a given prediction is as important as the prediction itself. When a message is flagged as harmful or dangerous, the XAI module provides a clear and human-readable

explanation allows users to understand which words, patterns, or contexts contributed to the decision. This transparency not only improves user trust but also offers an educational value by making users aware of the types of language or behavior that might be considered harmful in online communities. In turn, this can foster healthier digital interactions and reduce the risk of repeated offenses. From a cybersecurity perspective (Zhang et al., 2022), XAI is essential for ensuring that automated threat detection and anomaly detection systems remain auditable and reliable. Security analysts benefit from detailed, interpretable insights that can guide incident response and prioritization of alerts. For example, in a social platform monitoring scenario, XAI can help distinguish between false positives (benign messages incorrectly flagged as harmful) and true threats, improving system precision and reducing unnecessary intervention. Another significant benefit of XAI in these scenarios lies in regulatory and ethical compliance (Nannini, 2025). Many data protection frameworks emphasize the right to explanation when automated decision-making affects individuals. Integrating XAI into cyberbullying and cybersecurity tools ensures alignment with these principles and supports the development of responsible AI systems.

Following the classification and explainability stages, BPMN is leveraged to formally represent intervention strategies as process models, enabling traceability from detection events to preventive actions while embedding governance constraints such as privacy controls and accountability checkpoints.

3 Modeling of Cyberbullying Guidelines

In the context of cyberbullying detection and prevention, BPMN serves as a formalization layer that captures intervention logic, providing actionable support to end users. This design enables a clear separation between analytical processing and guideline modelling while ensuring continuity from detection to prevention. Among the wide range of policies and recommendations available, the following three key documents have emerged as the most relevant and have been adopted as reference sources for the BPMN-based modelling process:

- First, the international guidance provided by *UNICEF*[1] frames cyberbullying prevention as an integral component of child protection and digital rights. This perspective emphasizes the need for education on safe online practices, the empowerment of children and adolescents, capacity-building for parents and teachers, and stronger accountability mechanisms for digital platforms. UNICEF's approach highlights the importance of a multi-stakeholder ecosystem in safeguarding minors online.
- Second, the academic and practical guidance from the *Cyberbullying Research Center* (CRC) provides a set of evidence-based interventions that schools and local communities can adopt (Hinduja and Patchin, 2024). These include systematic assessment of bullying behaviors, the development of clear policies, specialized training for educators, continuous monitoring of incidents, and structured restorative responses. This framework is grounded in empirical research and focuses on practical measures that institutions can implement directly.
- Finally, the "Guidelines for the prevention and contrast of the phenomena of bullying and cyberbullying"[2] issued by the *Italian Ministry of Education and Merit* provide an operational framework tailored to the national school system. These guidelines translate broader principles into concrete responsibilities for schools, including defining school-level duties, establishing standardized reporting flows, mandating educational curricula on digital citizenship, and enhancing coordination with families, law enforcement, and local services. This national framework ensures consistency and accountability in the prevention and management of cyberbullying incidents across Italian educational institutions.

Overall, as reported in Table 1, the review of these three frameworks highlights a shared emphasis on education, early detection, coordinated intervention, and systemic collaboration as essential components of cyberbullying prevention and mitigation. Moreover, there is consensus on the necessity of multi-level interventions combining education, institutional policy, and technological safeguards. Early intervention and structured reporting mechanisms are critical to effective responses. Nonetheless, gaps remain, particularly regarding the accountability of private digital platforms. While UNICEF and research advocates emphasize platform responsibility, national school-centered frameworks often lack enforcement mechanisms beyond the school context. Additionally, the adoption of restorative practices remains uneven, and resource constraints pose practical barriers to effective program implementation.

[1] *https://www.unicef.org/stories/how-to-stop-cyberbullying*
[2] *https://www.mim.gov.it/web/guest/linee-guida-prevenzione-e-contrasto*

Table 1 - Comparison of guidelines for cyberbullying prevention and regulation

Dimension / Guidelines	**UNICEF**	**Cyberbullying Research Center**	**Italian Ministry of Education & Merit**
Primary Focus	Child protection, digital rights, online safety education	Evidence-based identification, prevention, school/community interventions	School policies, administrative procedures, prevention & contrast of cyberbullying
Scope	International, child-centered	International (research), school-centric but applicable to communities	National (Italy), mandatory for schools
Underlying Principles	Protection of child dignity, empowerment, multi-stakeholder collaboration	Evidence-driven, whole-school approach, assessment & monitoring	Prevention, early intervention, coordination with families and authorities
Key Stakeholders	Children, parents, educators, governments, digital platforms	Students, educators, school administrators, law enforcement	School leadership, teachers, students, families, regional authorities
Definition & Framing of Cyberbullying	Online violence and digital-harm violating children's rights	Repeated, intentional harm via digital devices; aligned with bullying research	Behaviors that cause harm through digital means; aligned with national legal framework
Prevention Strategies	Digital citizenship, awareness campaigns, safe online behavior, resilience-building	Policies, curriculum integration, staff training, parental involvement	Mandatory school plans, education programs, digital citizenship, teacher training
Intervention & Response	Encourage reporting, ensure psychosocial support, platform reporting tools	Multi-tiered interventions: prevention, targeted support, disciplinary and restorativ e actions	Structured procedures for identification, reporting, involvement of families, referral to authorities
Monitoring & Data	Promote research globally, encourage governments to track incidents	Strong emphasis on systematic assessment and evaluation	Schools must monitor, document, and evaluate actions regionally and nationally
Platform Accountability	Strong emphasis on platform safety mechanisms, improved reporting an d removal tools	Limited (more research-oriented than regulatory)	Minor emphasis; focus remains o n school-level actions
Legal/Regulator y Basis	Non-binding international guidance	Research-driven, non-regulatory recommendations	Italian law, Ministerial Decree n.18/2021
Strengths	Holistic, rights-based, globally adaptable	Evidence-based, deeply practical for educators	Operational, enforceable, tied to national education system
Limitations	High-level; implementation varies across countries	Requires school capacity and resources to implement effectively	Limited influence over social platforms; uneven implementation across regions

Although existing guidelines are often very specific and detailed, their structure and presentation tend to be highly verbose, legalistic or technical. This level of complexity can pose a significant barrier for non-expert users (e.g., parents, students, educators not specialized in law or cyber-safety) who need to interpret and apply them in concrete circumstances. Dense paragraphs of legal or procedural text may end up discouraging engagement, hindering clarity about concrete responsibilities, workflows, and rights, and limiting the actual implementation of the guidelines in real-world contexts. Adopting a visual modelling or informational design approach can significantly enhance comprehension and practical application. When applied to cyber-social guidelines (e.g., rules, reporting procedures, roles and responsibilities, intervention flows), such a visual modeling approach can serve several functions. First, it can lower the barrier to comprehension; even users unfamiliar with legal terminology or institutional jargon can follow a structured flowchart. In addition, visual models foster shared mental models among diverse stakeholders. Such shared understanding becomes especially valuable in urgent situations or where multiple actors must coordinate also enhancing scalability and reuse. Once procedures are

modelled, they can be reused, updated, and disseminated more easily, supporting training, awareness campaigns, and easier onboarding of new staff or community members.

Among current modelling approaches, including UML activity diagrams, flowcharts, data-flow models, and state-based representations, BPMN emerged as the most suitable framework for describing processes related to cyberbullying detection and support. Unlike simpler diagramming techniques, BPMN incorporates explicit semantics for events, decisions and roles, making it particularly well-suited for contexts that require coordination among several actors (e.g., educators, administrators, and families). As described in Zhou et al. (2018), BPMN can be useful for structuring and standardizing processes for cybersecurity detection, reporting, and support. By providing a graphical and formal representation of workflows, BPMN enables organizations to model how information flows, how responsibilities are distributed, and how interventions should be coordinated across different actors. Particularly, in the context of cyberbullying detection, BPMN models can describe processes such as incident identification, automated alert generation, data validation, escalation pathways, and communication with responsible personnel.

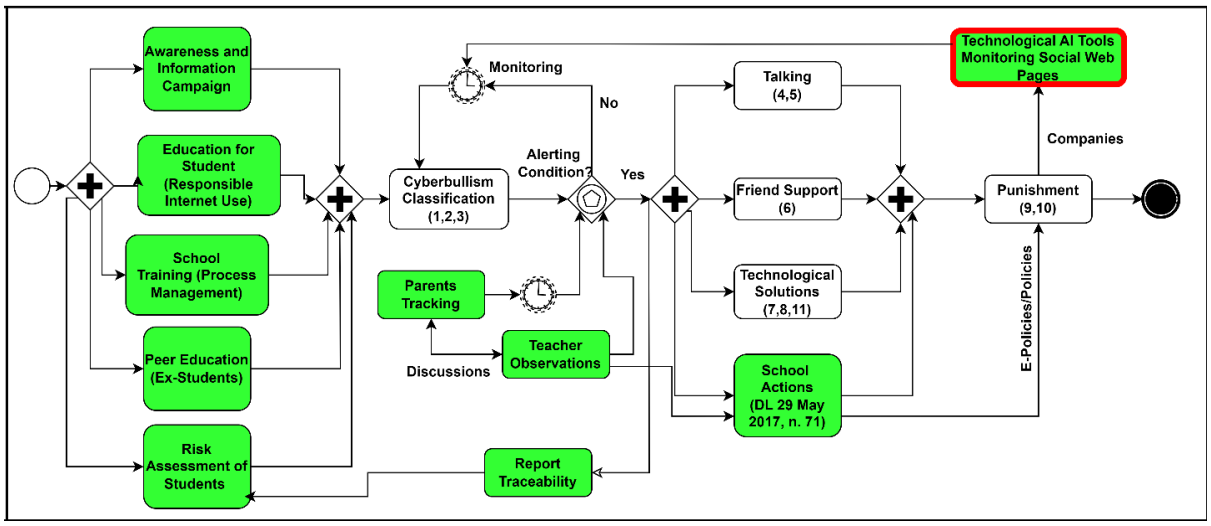


Figure 2 - BPMN model of national guidelines proposed by Italian Ministry of Education and Merit

The BPMN workflow may integrate automated monitoring tools (e.g., keyword detection or anomaly detection algorithms) with manual verification steps performed by teachers or designated school staff. Similarly, for support and response, BPMN can formally map actions including victim assistance, communication with families, documentation procedures, involvement of mental-health professionals, and, when necessary, referral to law enforcement or specialized services. For example, Figure 2 illustrates the processes defined by the Italian Ministry of Education and Merit. This BPMN diagram operationalizes the structured procedures mandated for Italian schools, including incident reporting, verification by designated school personnel, communication with families, activation of support measures, and coordination with local institutions or law-enforcement authorities when required. The model reflects the Ministry's school-centric approach by clearly delineating roles, responsibilities, and decision points that guide the response to cyberbullying cases within the national educational system.

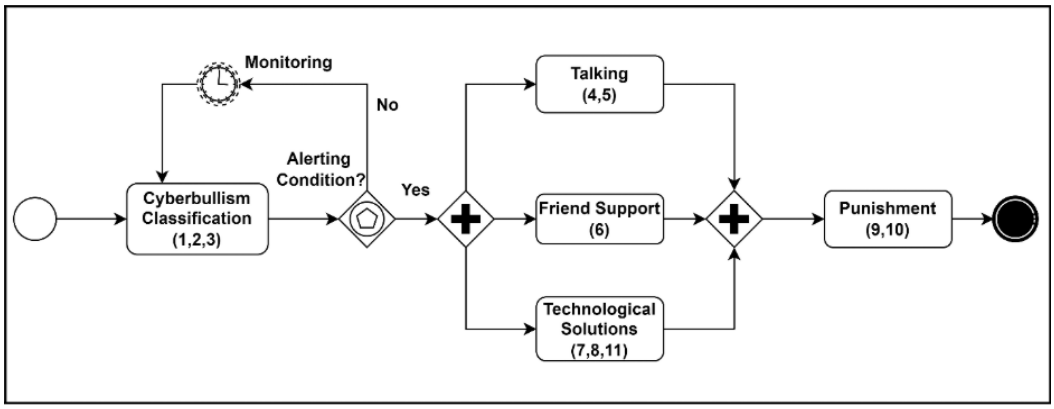


Figure 3 - BPMN model of UNICEF guidelines

Similarly, Figure 3 models the guidelines proposed by UNICEF, highlighting the multi-stakeholder nature of cyberbullying prevention as framed within child-protection and digital-rights principles. The BPMN model captures emphasis on awareness-raising, capacity-building for parents and educators, empowerment of children, and collaboration with digital platforms to enable effective reporting and response mechanisms. Events and message flows are used to illustrate how preventive education, early detection, and support actions interact across different actors.

However, the adoption of BPMN for cyberbullying detection and support also presents challenges. One limitation is that cyberbullying incidents are often context-dependent and emotionally sensitive, making it difficult to fully capture nuanced human judgment within a predefined process model. Additionally, effective workflows depend on accurate data collection and cooperation between multiple actors, e.g., students, teachers, administrators, families, and external agencies, which may not always be consistent or reliable. There are also technical challenges related to integrating BPMN-driven processes with automated detection tools, especially when these rely on privacy-sensitive data or advanced AI models. Finally, ensuring that all personnel are trained to understand and implement BPMN processes can require significant organizational effort. Despite these challenges, BPMN represents a valuable tool for operationalizing cyberbullying prevention and response efforts. By combining process transparency, standardization, and the ability to integrate both human and technological components, BPMN can help institutions create more effective, accountable, and timely interventions to protect young people in digital environments.

4 Conclusion and Future Work

In this work, we addressed the critical gap between the detection of cyberbullying incidents and the operationalization of effective prevention strategies. While AI-based approaches have significantly advanced the automatic identification of harmful content, their impact remains limited without structured mechanisms to guide consistent and timely interventions. To overcome this limitation, we proposed a unified framework that bridges detection and action by integrating AI-driven cyberbullying detection with process-oriented modeling through BPMN. First, we introduced an integrated architecture that combines AI-based detection capabilities with BPMN to enable traceable, structured, and auditable prevention processes. This integration allows detection outputs to seamlessly trigger predefined workflows, ensuring that responses to cyberbullying incidents are not only automated but also aligned with organizational policies and best practices. Second, we demonstrated how high-level prevention guidelines, often expressed as unstructured textual documents, can be systematically transformed into executable process workflows. This formalization enhances transparency by making decision logic explicit, improves accountability by clearly defining roles and responsibilities, and increases adaptability by allowing processes to be refined and extended as new requirements emerge. Moreover, the proposed approach lays a solid foundation for embedding risk management strategies and privacy-aware controls within AI-driven pipelines, supporting compliance with ethical and regulatory standards.

Future work will focus on extending the proposed framework along several directions. We aim to explore the integration of Generative AI techniques to enhance user interaction with prevention guidelines and process workflows. This includes the development of text-based, interactive explanation systems powered by large language models, capable of providing contextualized guidance, clarifying decision paths, and supporting stakeholders in navigating complex prevention procedures. Additionally, further research will investigate the incorporation of adaptive and self-improving workflows, as well as the evaluation of the framework in real-world scenarios to assess its effectiveness, usability, and scalability across different application domains.

5 Acknowledgement

This work has been supported by the project titled “Cyber Social Security” (CSS) - National Recovery and Resilience Plan, Mission 4 “Education and Research” – Component 2 “From Research to Business” – Investment 1.3, funded by the European Union – NextGenerationEU – CUP F53C22000740007.

6 Ethics Declaration

This research did not involve the use of personal data. Ethical clearance was not required.

7 AI Declaration

No AI tools were used in the writing of the paper.

References

Giumetti, G. W., and Kowalski, R. M. (2022). *Cyberbullying via social media and well-being*. Current opinion in psychology, 45, 101314.

Hinduja, S., and Patchin, J. W. (2024) *Cyberbullying: Identification, Prevention, & Response*. Cyberbullying Research Center.

Loseto, G., Gramegna, F., Pinto, A. et al. (2026). *Evaluating Trustworthiness and Interpretability of eXplainable AI Techniques for Cyberbullying Detection*. 1st International Workshop on Symbiotic AI for Cyber Social Security.

Nannini, L. (2025). *XAI: regulatory and ethical aspects*. In Explainable AI for Communications and Networking (pp. 181-207). Academic Press.

Rahman, M. M., Hossain, S., Bhusal, B. et al. (2025). *Cyber AI Trends: Future Trends in AI for Cyberbullying Prevention*. In Combating Cyberbullying With Generative AI (pp. 279-298). IGI Global Scientific Publishing.

Rosa, A., Massaro, A., Loseto, G., Santarsiero, F., and Schiuma, G. (2026). *Improving the organizational models: process mining as a tool for telediabetology hospital integration* (pp. 1-19). Business Process Management Journal.

Zhang, Z., Al Hamadi, H., Damiani, E., et al. (2022). *Explainable Artificial Intelligence Applications in Cyber Security: State-of-the-Art in Research*. IEEE Access, 10, 93104-93139.

Zhou, B., Maines, C., Tang, S. et al. (2018). *A Framework for the Visualisation of Cyber Security Requirements and its Application in BPMN*. In Guide to Vulnerability Analysis for Computer Networks and Systems: An Artificial Intelligence Approach (pp. 339-366). Cham: Springer International Publishing.

Research Methods for Sustainable Organizations Using Intelligent Knowledge

Pierluigi Mascaro
Mercatorum Uniuversity, Rome, Italy

Abstract

The transition toward sustainable organizations has intensified the need for advanced knowledge systems capable of addressing complex economic, environmental, and social challenges. In this context, intelligent knowledge—defined as the integration of human expertise, organizational learning, and digital technologies such as artificial intelligence (AI) and data analytics—has emerged as a critical enabler of informed decision-making and long-term organizational resilience. This paper examines the research methods most suitable for investigating sustainable organizations through the lens of intelligent knowledge, with particular attention to their ability to capture complexity, dynamism, and socio-technical interactions.

The study adopts an integrative methodological perspective, combining insights from quantitative, qualitative, mixed-methods, design science, and participatory research approaches. Quantitative methods support the analysis of measurable relationships between knowledge capabilities and sustainability performance, while qualitative approaches provide in-depth understanding of organizational processes, cultural dynamics, and meaning-making practices. Mixed-methods designs enable the triangulation of findings, enhancing both validity and explanatory power. In addition, design science and action research approaches contribute to bridging theory and practice by supporting the development and evaluation of intelligent knowledge systems in real organizational contexts.

A central argument of the paper is that no single methodological approach is sufficient to address the multifaceted nature of sustainability challenges. Instead, the research highlights the necessity of methodological pluralism, emphasizing longitudinal and process-oriented designs, multi-level analysis, and the integration of ethical considerations related to AI and data use. The paper also underscores the importance of reflexivity and interdisciplinary collaboration in developing robust and context-sensitive research frameworks.

By providing a comprehensive overview of methodological approaches and their implications, this study contributes to the advancement of sustainability research and offers a coherent framework for analyzing how intelligent knowledge can support sustainable organizational transformation. The proposed methodological orientation is both theoretically grounded and practically relevant, enabling researchers and practitioners to better understand and navigate the complexities of sustainability in an increasingly data-driven world.

Keywords: Sustainable Organizations; Intelligent Knowledge; Research Methodologies; Artificial Intelligence; Socio-Technical Systems.

Paper type: Academic Research Paper

1 Introduction

The growing urgency of environmental degradation, climate change, and social inequality has fundamentally reshaped the role of organizations in contemporary society. Organizations are no longer evaluated solely on their economic performance but increasingly on their ability to generate sustainable value across environmental and social dimensions. This shift has led to the emergence of the concept of sustainable organizations, which integrate economic, environmental, and social objectives into their strategies and operations (Elkington, 1997; Dyllick & Muff, 2016).

At the same time, rapid advancements in digital technologies have transformed how knowledge is created, shared, and applied within organizations. Artificial intelligence, big data analytics, and digital platforms have enabled new forms of intelligent knowledge—knowledge that is not only accumulated but also processed, interpreted, and applied through advanced computational systems (Brynjolfsson & McAfee, 2017; Russell & Norvig, 2021). This convergence of sustainability and intelligent knowledge presents both opportunities and challenges for organizations seeking to navigate complex and uncertain environments.

Despite the growing importance of intelligent knowledge in sustainability, there remains a need for robust research methodologies capable of capturing its multidimensional nature. Traditional research approaches often fall short in addressing the dynamic, socio-technical, and multi-level characteristics of sustainable organizations. This paper aims to fill this gap by examining research methods suitable for studying sustainable organizations using intelligent knowledge and by analyzing the methodological implications of this research domain.

This transformation also implies a broader reconfiguration of how organizations conceptualize knowledge, shifting from linear and static models toward dynamic, adaptive, and systemically embedded understandings. In

such contexts, knowledge is increasingly shaped by the interaction between human cognition and algorithmic systems, requiring new analytical lenses capable of capturing co-evolutionary processes between technology, organization, and environment.

2 Theoretical Background

2.1 Sustainable organizations

The concept of sustainable organizations is rooted in the idea of the triple bottom line, which emphasizes the integration of economic, environmental, and social performance (Elkington, 1997). Over time, this concept has evolved into more comprehensive frameworks that highlight the systemic and transformative nature of sustainability in organizations (Dyllick & Muff, 2016).

Sustainable organizations are characterized by their ability to adapt to changing environmental conditions, manage stakeholder relationships, and innovate in ways that reduce negative impacts while creating long-term value. This requires not only technological capabilities but also organizational learning, cultural change, and strategic alignment (Hart & Dowell, 2011).

The systemic and multi-dimensional nature of sustainable organizations implies that their study requires multi-level and context-sensitive research designs. Methodologically, this entails combining firm-level analysis with broader ecosystem perspectives, including stakeholders, institutions, and environmental systems. It also necessitates the integration of qualitative approaches to capture cultural and strategic dimensions, alongside quantitative measures of sustainability performance. Furthermore, the adaptive and evolving character of sustainable organizations reinforces the importance of longitudinal and comparative studies, enabling researchers to observe how sustainability practices and outcomes develop over time across different organizational and institutional contexts.

2.2 Intelligent knowledge

Intelligent knowledge builds on the knowledge-based view of the firm, which considers knowledge as the most strategically significant resource (Grant, 1996). It extends this perspective by incorporating digital technologies and AI into knowledge processes. Intelligent knowledge involves the interaction between human cognition, organizational routines, and computational systems (Nonaka & Takeuchi, 1995; Davenport & Prusak, 1998).

The integration of AI and data analytics allows organizations to process large volumes of data, identify patterns, and support decision-making in real time. However, intelligent knowledge is not purely technological; it also depends on social and organizational contexts, including trust, culture, and governance (Orlikowski, 2007).

The hybrid nature of intelligent knowledge—situated at the intersection of human, organizational, and technological domains—requires the adoption of socio-technical and multi-method research designs. Methodologically, this implies combining computational techniques (e.g., data analytics, machine learning validation) with qualitative approaches capable of capturing human interpretation, organizational routines, and cultural dimensions. It also necessitates careful consideration of data governance, algorithmic transparency, and epistemological assumptions embedded in AI systems. As a result, researchers must integrate technical rigor with interpretive depth, ensuring that both the performance and the meaning of intelligent knowledge are adequately analyzed.

2.3 Socio-technical systems perspective

Understanding intelligent knowledge in sustainable organizations requires a socio-technical perspective that considers both technological and social elements. Socio-technical systems theory emphasizes the interdependence between human actors and technological systems, highlighting the importance of alignment between them (Checkland, 1999).

This perspective is particularly relevant in sustainability research, where technological solutions must be integrated with organizational practices and societal values. It also underscores the need for research methods that can capture these complex interactions (Geels, 2002).

Adopting a socio-technical systems perspective entails methodological approaches that explicitly account for interactions, interdependencies, and co-evolution between social and technological components. This requires the use of integrative and systemic research designs, such as case studies, system dynamics modeling, and network analysis, capable of capturing relational complexity. Furthermore, it implies that unit-of-analysis choices must extend beyond individual organizations to include systems, ecosystems, and transitions.

Consequently, researchers must embrace methodological pluralism and develop frameworks that can bridge micro-level practices with macro-level transformations, ensuring a holistic understanding of sustainability processes.

3 Research Methodologies

3.1 Quantitative methods

Quantitative research methods are essential for analyzing measurable relationships between intelligent knowledge and sustainability outcomes. Surveys and statistical models allow researchers to test hypotheses and identify patterns across large samples (Alavi & Leidner, 2001).

Structural equation modeling (SEM) is widely used to examine relationships between latent variables such as knowledge management capabilities and sustainability performance. Panel data analysis enables the study of changes over time, providing insights into causal relationships (Bansal & DesJardine, 2014).

Big data analytics and machine learning further enhance quantitative research by enabling the analysis of large and complex datasets. These methods are particularly useful for studying energy consumption, supply chain efficiency, and environmental impacts (Wamba et al., 2015). However, they also introduce challenges related to data quality, bias, and interpretability.

The increasing reliance on quantitative and data-driven approaches implies a need for methodological rigor in data management, model specification, and validation procedures. Researchers must address issues related to data quality, completeness, and representativeness, particularly when using large-scale or unstructured datasets. Moreover, the use of advanced analytical techniques such as machine learning requires careful consideration of model transparency, reproducibility, and interpretability. This leads to a growing emphasis on hybrid analytical frameworks that combine predictive accuracy with explanatory clarity, ensuring that quantitative findings remain both robust and meaningful in sustainability contexts.

3.2 Qualitative methods

Qualitative research methods provide in-depth insights into the processes and contexts through which intelligent knowledge is developed and applied. Case studies are particularly valuable for exploring complex phenomena in real-world settings (Yin, 2018).

Interviews and ethnographic studies allow researchers to understand how individuals and groups interpret and use knowledge in sustainability initiatives. These methods are essential for capturing organizational culture, leadership dynamics, and ethical considerations (Eisenhardt, 1989).

Longitudinal qualitative research is especially important for studying organizational change over time. It enables researchers to trace the evolution of knowledge practices and sustainability strategies (Hart & Dowell, 2011).

The use of qualitative methods requires strong attention to contextualization, researcher reflexivity, and interpretive validity. Researchers must ensure transparency in data collection and analysis procedures, including coding strategies and case selection criteria. Furthermore, the emphasis on process and meaning implies that findings are often context-specific, which necessitates careful consideration of transferability rather than generalizability. Longitudinal qualitative designs also introduce challenges related to access, consistency, and data continuity, reinforcing the need for well-structured research protocols and sustained engagement with field settings.

3.3 Mixed-methods approaches

Mixed-methods research combines quantitative and qualitative approaches to provide a more comprehensive understanding of complex phenomena. By integrating numerical data with contextual insights, researchers can enhance the validity and relevance of their findings (Creswell & Plano Clark, 2018).

Sequential designs involve conducting one type of research before the other, while concurrent designs integrate both simultaneously. Mixed-methods approaches are particularly useful in sustainability research, where multiple perspectives and levels of analysis are required (Saunders et al., 2019).

The adoption of mixed-methods approaches introduces the challenge of methodological integration, requiring researchers to ensure coherence between quantitative and qualitative components. This includes aligning research questions, data collection strategies, and analytical frameworks across methods. Additionally, mixed-methods designs demand careful consideration of timing, priority, and weighting of different data

sources. Successfully addressing these challenges enhances triangulation and strengthens the robustness of findings, but also requires advanced methodological expertise and clear justification of design choices.

3.4 Design science research

Design science research focuses on the creation and evaluation of artifacts, such as models, systems, and tools. It is particularly relevant for developing intelligent knowledge systems that support sustainability (Hevner et al., 2004).

This approach involves iterative cycles of design, implementation, and evaluation, allowing researchers to refine solutions based on feedback. Action design research further integrates practical intervention with theoretical development (Sein et al., 2011).

Design science research shifts the role of the researcher from passive observer to active designer, which introduces important implications for evaluation and rigor. Methodologically, this requires clearly defined criteria for artifact effectiveness, utility, and generalizability. It also necessitates iterative validation processes involving stakeholders and real-world testing environments. Furthermore, the dual aim of theoretical contribution and practical relevance demands transparent documentation of design choices and learning cycles, ensuring that results are both scientifically credible and practically applicable.

3.5 Participatory and action research

Participatory and action research methods involve stakeholders in the research process, promoting collaboration and co-creation of knowledge. These approaches are especially important in sustainability research, where multiple stakeholders are involved (Reason & Bradbury, 2008).

Action research emphasizes learning through action and reflection, enabling organizations to implement changes while generating knowledge. This approach bridges the gap between theory and practice (Davenport & Prusak, 1998).

Participatory and action research approaches require a redefinition of the researcher's role, emphasizing collaboration, facilitation, and ethical responsibility. Methodologically, this entails managing power dynamics, ensuring inclusivity, and maintaining transparency throughout the research process. It also introduces challenges related to objectivity and bias, as researchers become embedded within the context they study. To address these issues, reflexivity, documentation, and iterative validation are essential. At the same time, these approaches enhance the practical relevance and societal impact of research, aligning closely with the goals of sustainability-oriented inquiry.

4 Methodological Implications

The integration of intelligent knowledge into sustainability research has significant methodological implications. First, it requires a shift toward longitudinal and process-oriented research designs. Sustainability and knowledge processes evolve over time, making static analyses insufficient.

Second, it necessitates multi-level analysis. Intelligent knowledge operates at individual, organizational, and societal levels, requiring methods that can capture interactions across these levels.

Third, it introduces ethical considerations related to AI and data usage. Researchers must address issues such as bias, transparency, and accountability.

Fourth, it blurs the boundary between research and practice. Methods such as design science and action research involve active participation in organizational processes, requiring reflexivity and ethical awareness.

Finally, it reinforces the importance of methodological pluralism. No single method can capture the complexity of sustainable organizations using intelligent knowledge.

These implications collectively require the development of integrative methodological frameworks that explicitly address temporality, scale, and ethics within a unified research design. Researchers must adopt flexible and adaptive methodologies capable of evolving alongside the phenomena under investigation. This includes combining longitudinal data collection with multi-level modeling techniques, integrating ethical assessment into research protocols, and ensuring continuous reflexivity throughout the research process. As a result, methodological rigor is increasingly defined not only by technical precision but also by the ability to capture complexity, ensure ethical responsibility, and generate actionable insights.

Methodologically, this means moving beyond isolated techniques toward coordinated research designs that combine longitudinal observation, multi-level modeling, and reflexive inquiry. It also requires embedding ethical governance directly into the research process, particularly when dealing with AI-driven systems and data-

intensive environments, ensuring that methodological choices remain aligned with both scientific rigor and societal responsibility.

5 Discussion

The findings highlight the need for integrative research approaches that combine multiple methodologies. Quantitative methods provide generalizable insights, while qualitative methods offer contextual understanding. Mixed-methods approaches bridge these perspectives, and design science and action research contribute practical solutions.

The role of intelligent knowledge in sustainability is likely to grow as technologies continue to evolve. Future research should explore emerging technologies such as explainable AI and their implications for sustainability.

The discussion underscores the necessity of designing research strategies that are inherently hybrid and forward-looking. Methodologically, this implies not only combining existing approaches but also continuously adapting them to emerging technological contexts. Researchers must develop capabilities to study evolving phenomena such as explainable AI, which require new evaluation criteria, interdisciplinary knowledge, and innovative data collection techniques. This dynamic orientation reinforces the need for methodological flexibility and continuous methodological innovation as core competencies in sustainability research.

This discussion implies that future research agendas must increasingly prioritize adaptability and methodological responsiveness to technological evolution. In particular, the growing presence of intelligent and explainable AI systems requires researchers to continuously update analytical frameworks, evaluation criteria, and data interpretation strategies. Methodologically, this translates into a stronger emphasis on iterative research designs, cross-disciplinary integration, and continuous validation processes, ensuring that methodological tools remain aligned with rapidly changing socio-technical realities.

6 Conclusion

This paper has examined research methods for studying sustainable organizations using intelligent knowledge. It has highlighted the importance of combining different methodological approaches to address the complexity of sustainability challenges.

The integration of intelligent knowledge requires new research paradigms that are adaptive, interdisciplinary, and practice-oriented. By adopting such approaches, researchers can contribute to both theoretical advancement and practical impact.

The conclusions of this study emphasize that future research must move beyond methodological fragmentation toward coherent and integrative research paradigms. This entails developing unified frameworks that systematically combine different methods while maintaining clarity in research design and purpose. Additionally, it highlights the importance of aligning methodological choices with both theoretical objectives and practical relevance, ensuring that research outcomes effectively support sustainable transformation. Ultimately, methodological innovation becomes a central driver of progress in the study of sustainable organizations using intelligent knowledge.

The enrichment of the conclusion reinforces the need for a holistic and forward-looking methodological orientation. Researchers must design studies that are not only methodologically sound but also capable of capturing emerging phenomena, integrating ethical considerations, and producing scalable insights. This entails combining rigor with flexibility, embracing interdisciplinary collaboration, and fostering continuous methodological learning. Ultimately, the ability to effectively study sustainable organizations using intelligent knowledge will depend on the development of research frameworks that are dynamic, integrative, and closely aligned with the evolving realities of both organizations and the broader socio-technical systems in which they operate.

7 Ethical Declarations

This study is based exclusively on theoretical and secondary literature analysis and does not involve human participants, clinical data, or any form of experimental intervention. As such, no formal ethical approval was required. Nevertheless, the research adheres to established academic standards of integrity, transparency, and responsible scholarship. All sources have been appropriately cited to ensure intellectual honesty and to avoid plagiarism. Particular attention has been given to the ethical implications of artificial intelligence and data-driven systems, especially regarding issues of transparency, bias, and accountability in the context of sustainable organizational research.

8 AI Declaration

Artificial intelligence tools were used in this study solely as a linguistic and editorial support instrument to assist with text refinement, structural organization, and clarity of expression. All conceptual development, analytical framing, and scientific content selection were conducted under full human authorship and responsibility. The use of AI did not involve autonomous generation of research data, empirical analysis, or scientific inference. The authors confirm that they critically reviewed, edited, and validated all AI-assisted outputs to ensure academic accuracy, coherence, and compliance with research integrity standards.

References

Alavi, M., & Leidner, D. E. (2001). Knowledge management and knowledge management systems. *MIS Quarterly*, 25(1), 107–136.

Bansal, P., & DesJardine, M. R. (2014). Business sustainability. *Strategic Organization*, 12(1), 70–78.

Brynjolfsson, E., & McAfee, A. (2017). *Machine, Platform, Crowd*. Norton.

Checkland, P. (1999). *Systems Thinking, Systems Practice*. Wiley.

Creswell, J. W., & Plano Clark, V. L. (2018). *Mixed Methods Research*. SAGE.

Davenport, T. H., & Prusak, L. (1998). *Working Knowledge*. Harvard.

Dyllick, T., & Muff, K. (2016). Sustainable business. *Organization & Environment*, 29(2), 156–174.

Eisenhardt, K. M. (1989). Case study research. *Academy of Management Review*, 14(4), 532–550.

Elkington, J. (1997). *Cannibals with Forks*. Capstone.

Geels, F. W. (2002). Technological transitions. *Research Policy*, 31(8–9), 1257–1274.

Grant, R. M. (1996). Knowledge-based theory. *Strategic Management Journal*, 17(S2), 109–122.

Hart, S. L., & Dowell, G. (2011). Natural-resource-based view. *Journal of Management*, 37(5), 1464–1479.

Hevner, A. R., et al. (2004). Design science research. *MIS Quarterly*, 28(1), 75–105.

Nonaka, I., & Takeuchi, H. (1995). *The Knowledge-Creating Company*. Oxford.

Orlikowski, W. J. (2007). Sociomaterial practices. *Organization Studies*, 28(9), 1435–1448.

Reason, P., & Bradbury, H. (2008). *Action Research*. SAGE.

Russell, S., & Norvig, P. (2021). *Artificial Intelligence*. Pearson.

Saunders, M., et al. (2019). *Research Methods for Business Students*. Pearson.

Sein, M. K., et al. (2011). Action design research. *MIS Quarterly*, 35(1), 37–56.

Wamba, S. F., et al. (2015). Big data analytics. *Journal of Business Research*, 70, 356–365.

Yin, R. K. (2018). *Case Study Research*. SAGE.

Research on Megatrends and the Evolution of ESG Standard

Giovanni Maiuri[2], Domenico Celenza[1], Alessia Nardecchia[2], Francesco Di Vona[2]
[1] University of Cassino and Southern Lazio, Cassino, Italy
[2] Independent scholar

Abstract

This paper examines how global megatrends can be transformed into actionable knowledge for the design of sustainability standards and the development of predictive environmental, social and governance reporting. The argument is grounded in the view that intelligent and sustainable organizations are not merely adopters of digital tools, but knowledge-processing systems that detect weak signals, translate them into governance priorities and combine human and artificial intelligence to support responsible value creation. The paper develops a conceptual framework linking megatrend research, European sustainability reporting traditions, knowledge management and artificial intelligence. The analysis identifies three historically rooted European trajectories: the German emphasis on measurement, comparability and procedural discipline; the French emphasis on public regulation, social responsibility and corporate accountability; and the Italian "Economia Aziendale" tradition, which connects economic performance, social embeddedness and systemic equilibrium. The paper argues that the next stage in the evolution of sustainability standards requires a shift from descriptive disclosure to predictive ESG intelligence, understood as the capacity to transform megatrend knowledge into forward-looking, scenario-based and accountable sustainability decisions. Artificial intelligence is interpreted as a cognitive infrastructure that can strengthen human interpretation by mapping weak signals, clustering heterogeneous information, modelling scenario-based risks and connecting sustainability variables with business continuity and organizational resilience. The contribution of the paper is twofold. First, it reframes ESG standards as dynamic knowledge artefacts rather than static compliance devices. Second, it proposes an AI-enabled pathway through which organizations can move from fragmented non-financial reporting to anticipatory, accountable and human-centred sustainability governance. An illustrative climate-transition example shows how a megatrend can be translated into dynamic ESG metrics, predictive outputs and strategic governance decisions. The paper concludes that predictive ESG intelligence can support intelligent organizations and communities only when embedded in ethical oversight, stakeholder dialogue and a relational view of value creation.

Keywords: megatrends; ESG standards; knowledge management; predictive analytics; artificial intelligence

Paper type: Academic Research Paper

1 Introduction

The future of organizations and communities increasingly depends on their capacity to combine knowledge, intelligence and sustainability. Demographic change, technological acceleration, climate pressure, geopolitical instability and evolving social expectations no longer operate as distant variables. They enter strategy, reshape business-model legitimacy and influence how value creation is evaluated by investors, regulators, workers, customers and communities. The ability to detect, interpret and govern megatrends has therefore become a condition of sustainable resilience (Ansoff, 1975). This paper starts from the proposition that sustainability standards are not neutral technical instruments. They are knowledge structures through which societies define what should be observed, measured, compared and reported. ESG standards reflect a changing understanding of the firm and of its relationship with the wider ecosystem. When reporting systems move beyond financial data, they recognise that organizational continuity depends on social trust, ecological viability, institutional accountability and stakeholder relations. From a knowledge management perspective, reporting identifies, codifies, shares and uses relevant knowledge to support decisions (Rohrbeck and Kum, 2018).

The debate on intelligent organizations and communities adds a further dimension. Intelligence cannot be reduced to artificial intelligence or individual cognition. It includes collective sensemaking, organizational learning, community participation and the ability to connect dispersed information with coherent action. Intelligent organizations are hybrid systems in which human judgement, digital infrastructures and collaborative practices interact. Technologies expand learning capacity, but organizations remain accountable for how knowledge is selected, interpreted and applied.

The aim of the paper is to develop a conceptual framework on megatrends and ESG standards, aligned with the IFKAD debate on intelligent knowledge for sustainable organizations. The research question is: how can organizations transform megatrend analysis into predictive ESG intelligence for sustainable value creation and resilience? The answer is theoretical and integrative, combining European corporate social responsibility, knowledge-based views of the firm and AI-enabled predictive analytics. Predictive ESG intelligence denotes a

knowledge-based dynamic capability through which an organization transforms megatrend knowledge into anticipatory evaluations for strategy, risk management and resilience. Cognitive infrastructure is the socio-technical framework through which artificial intelligence structures information flows and helps human agents identify patterns, anomalies and emerging risks. Dynamic ESG metrics preserve comparability while remaining responsive to new risks and regulatory changes. Knowledge artefacts are standards, metrics, reports or models that allow organizations to codify and disseminate knowledge. ESG standards can therefore define how sustainability phenomena are observed, measured and discussed in governance processes.

2 Theoretical Background

2.1. Megatrends as strategic knowledge

Megatrends are long-term transformations that reshape the conditions under which organizations operate. They include demographic ageing, urbanization, digitalization, climate transition, resource scarcity, geopolitical fragmentation and changing expectations concerning work and quality of life. Their relevance lies not only in their scale, but also in their systemic character. A megatrend rarely affects one function or industry. It alters markets, institutions, technologies and social practices simultaneously, so organizations that treat it as an isolated risk tend to underestimate its cumulative effects.

Megatrends can be understood as strategic knowledge objects because they are not immediately available as stable facts. They must be constructed through data collection, interpretation, comparison and dialogue. Organizations decide which signals are meaningful, which sources are credible and which time horizons matter. The knowledge-based theory of the firm suggests that advantages depend on the capacity to integrate specialized knowledge (Grant, 1996). Megatrend analysis extends this view by requiring organizations to integrate external, uncertain and ambiguous knowledge (Weick, 1995).

This requirement recalls Nonaka and Takeuchi (1995), who described knowledge creation as a dynamic interaction between tacit and explicit knowledge. Sustainability challenges require this interaction. Quantitative indicators are necessary, but they do not exhaust the meaning of climate vulnerability, community well-being or organizational resilience. Tacit knowledge from workers, managers, communities and stakeholders helps interpret data and prevents narrow measurement. Intelligent ESG governance therefore requires mechanisms through which data become shared understanding.

The treatment of megatrends as knowledge also connects sustainability to resource-based and dynamic capability perspectives. Knowledge becomes strategic when it is difficult to imitate, embedded in routines and recombined with external signals (Barney, 1991; Teece et al., 1997). In digital transformation, this recombination depends on data quality, interpretive routines and capabilities for translating information into action (Davenport and Prusak, 1998; De Mauro et al., 2016; Vial, 2019). Megatrends such as climate transition or demographic ageing should therefore be read as socio-technical transitions requiring distributed knowledge, experimentation and coordinated learning (Geels, 2002; George et al., 2016).

2.2. Sustainability standards as institutionalised knowledge

Sustainability standards identify disclosure categories, define metrics, impose comparability and support accountability. They also shape internal learning. When an issue enters a reporting framework, it becomes more visible to managers and governance bodies. What is measured tends to enter routines, risk maps and performance discussions. ESG standards are therefore institutionalised knowledge: they stabilize social expectations and translate them into organizational language.

The rise of non-financial reporting reflects a broader transformation in the concept of the firm. Classical financial reporting privileges profitability, solvency and asset valuation. Sustainability reporting expands the field by including conditions that make value creation durable and legitimate. This expansion is consistent with European corporate social responsibility and with approaches that view the enterprise as an open system. In the Italian Economia Aziendale tradition, the firm is a coordinated institution whose equilibrium depends on economic, social and environmental relations (Ferrero, 1968; Onida, 1971; Borgonovi and Anessi Pessina, 2018).

The contemporary European agenda can therefore be read as regulatory innovation and intellectual continuity. The Corporate Sustainability Reporting Directive reinforced the idea that sustainability information is relevant for governance and stakeholders, not merely reputation (European Commission, 2022). Yet the underlying principle is older: the enterprise must be understood in relation to the community and to the conditions of continuity. Cassandro (1978) expressed this idea by presenting the firm as a responsible actor whose conduct has social consequences.

Sustainability also creates tensions. Organizations must balance short-term performance with long-term resilience, shareholder expectations with stakeholder claims and standardized indicators with context-specific meanings. Such tensions are intrinsic to sustainability and should not be eliminated prematurely (Hahn et al., 2010; Van der Byl and Slawinski, 2015). ESG standards can support this balance only if treated as learning devices, not only as compliance checklists.

This institutionalization is visible in European and global reporting architectures. The European Sustainability Reporting Standards operationalize double materiality, while International Sustainability Standards Board standards emphasize decision-useful sustainability-related financial information (European Commission, 2023; EFRAG, 2023; IFRS Foundation, 2023a; IFRS Foundation, 2023b). GRI, ISO 26000 and TCFD contributed to a shared vocabulary for responsibility, impact and risk (GRI, 2021; ISO, 2010; Task Force on Climate-related Financial Disclosures, 2017). These developments extend stakeholder theory and sustainable development thinking, from the Brundtland definition to triple bottom line debates (Freeman, 1984; World Commission on Environment and Development, 1987; Elkington, 1997).

2.3. Artificial intelligence as cognitive infrastructure

Artificial intelligence introduces new possibilities for sustainability reporting because it can process large and heterogeneous datasets beyond traditional managerial routines. Machine learning, natural language processing and predictive analytics may help organizations analyse public documents, regulatory changes, scientific evidence, market signals, stakeholder sentiment and operational data. These tools detect patterns, identify anomalies and model alternative scenarios, extending the statistical and computational capacity for inference and prediction (Hastie et al., 2009; Russell and Norvig, 2021).

AI should not, however, be framed as an autonomous solution. The central issue is not only prediction, but interpretation. Predictive systems may estimate probabilities, but governance bodies decide which risks are material, which trade-offs are acceptable and which actions are legitimate. This distinction is important in ESG reporting, where decisions affect diverse stakeholders, communities, employees and future generations. AI strengthens organizational intelligence only when embedded in human judgement, transparency and accountability.

The concept of cognitive infrastructure is useful here. AI systems can organize information flows, support sensemaking and make weak signals more visible. They can also connect sustainability variables with financial and operational risk. However, they remain part of a broader socio-technical system. If data are biased or poorly interpreted, outputs may reinforce narrow assumptions. If governance processes are weak, predictive analytics may become decorative. Predictive ESG intelligence therefore requires technological capability and organizational maturity.

3 European Trajectories in the Evolution of ESG Standards

Three European trajectories are useful for understanding the evolution of sustainability standards: German, French and Italian. They should not be treated as rigid national stereotypes, but as historically formed orientations that emphasize different dimensions of organizational knowledge. Their convergence helps explain why European sustainability reporting combines measurement, regulation, narrative disclosure and systemic thinking.

The German trajectory is associated with structured indicators, measurability, standardization and procedural discipline. Sustainability information gains credibility when supported by precise metrics and comparable methods. Research on German non-financial reporting shows the movement from voluntary disclosure to mandatory obligations after the CSR Directive Implementation Act, while emphasizing report quality and links between financial and non-financial information (Hoffmann, Dietsche and Hobelsberger, 2018). The German Sustainability Code reflects the same orientation through 20 criteria for transparency and comparability (German Council for Sustainable Development, 2020). This contribution strengthens the reliability of sustainability standards and supports AI-enabled reporting, because predictive analytics requires structured data.

The French trajectory assigns a strong role to public regulation, corporate accountability and social responsibility. It emphasizes institutions in defining what organizations owe to society. This legal tradition is visible in the Nouvelles Régulations Économiques framework and in later reforms that expanded social and environmental disclosure. Studies of French CSR disclosure confirm that national legal, cultural and institutional traditions shape the legitimacy of corporate responsibility reporting (Antal and Sobczak, 2007; Chauvey et al., 2015). ESG standards therefore express public expectations concerning labour, inclusion, governance quality, environmental transition and corporate legitimacy.

The Italian trajectory, rooted in Economia Aziendale, offers a systemic interpretation (Zappa, 1927). The firm is viewed as an institution that must preserve equilibrium across economic, patrimonial, social and organizational dimensions. Onida's conception of the enterprise as an open system implies that continuity cannot be reduced to financial solvency (Onida, 1971). Cassandro's institutional view reinforces the firm's responsibilities towards the broader community (Cassandro, 1978). This perspective connects sustainability with business continuity, stakeholder relations and long-term viability. Taken together, these trajectories provide a European basis for moving from descriptive reporting to anticipatory governance. The German emphasis on measurement supports data quality. The French emphasis on accountability supports legitimacy. The Italian emphasis on systemic equilibrium supports integrated interpretation. Predictive ESG intelligence requires all three dimensions: measurable data, public responsibility and a systemic understanding of value creation.

4 An AI-Enabled Pathway to Predictive ESG Intelligence

The proposed framework defines predictive ESG intelligence as the capacity to transform megatrend knowledge into forward-looking sustainability decisions through human judgement, collective learning and AI-enabled analytics. This capacity develops through three phases: mapping megatrends and weak signals, embedding them into dynamic ESG metrics and connecting them with scenario-based sustainability risk intelligence. In practice, organizations move among these phases as new information emerges.

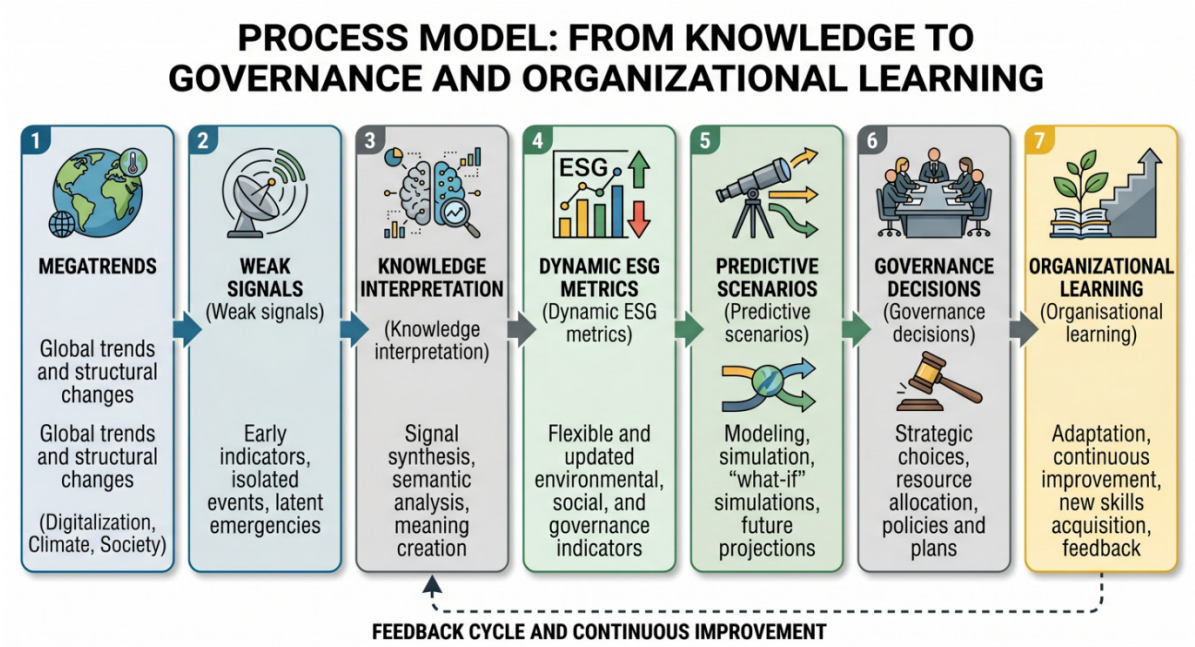


Figure 1. From megatrends to predictive ESG intelligence.

The figure shows how organizations can transform broad external transformations into actionable sustainability knowledge. AI supports the detection and organization of weak signals, while human and collective judgement remain responsible for interpretation, materiality assessment and governance action.

4.1. Mapping megatrends and weak signals

The first phase concerns the identification and interpretation of megatrends. Organizations collect large volumes of information, but much remains fragmented across departments, platforms and external sources. AI can assist through semantic analysis, text mining and thematic clustering, organizing heterogeneous material into interpretable domains. Climate regulation, energy prices, technological disruption and supply-chain instability may be connected as broader transition risk. The value of AI lies in making such relationships more visible.

Mapping is not a purely technical act. It requires the definition of boundaries, categories and relevance criteria. A weak signal may be insignificant in one context and strategically important in another. Organizations must therefore combine automated detection with deliberative interpretation. Cross-functional teams, sustainability experts, risk managers and stakeholders can help evaluate whether a trend is material. This

collaborative process prevents predictive systems from becoming detached from organizational knowledge and local experience.

The output of this phase is a structured megatrend knowledge base. It should be updated as new evidence, regulations and stakeholder expectations emerge, and it should distinguish between transient noise and persistent transformation. This distinction matters because sustainability governance requires attention to long-term change, whereas digital information environments often reward immediate reaction.

4.2. Embedding megatrends into dynamic ESG metrics

The second phase translates megatrend knowledge into ESG metrics and reporting categories. Traditional disclosure systems often describe past performance. Predictive ESG intelligence adds forward-looking indicators showing how external transformations may affect future conditions, including climate exposure, dependence on scarce resources, workforce skills, social acceptance of business models and governance readiness for regulation.

Dynamic metrics do not abandon standardization. They require a stable architecture for comparison over time and across units, while remaining flexible enough to incorporate emerging risks. A reporting system that cannot register new social or environmental vulnerabilities becomes obsolete. The challenge is to combine comparability with adaptability through processes for revising indicators, documenting assumptions and explaining why measures change.

The incorporation of megatrends into ESG metrics also changes the meaning of materiality. Materiality is not only a backward-looking assessment of issues that have already affected performance. It becomes an anticipatory judgement about issues that may influence value creation, stakeholder well-being and organizational continuity. Predictive analytics can support this judgement, but materiality remains a governance decision involving values, responsibilities and accountability.

4.3. Developing scenario-based sustainability risk intelligence

The third phase connects ESG indicators with scenario-based risk intelligence. Predictive systems can model how different megatrends may combine under alternative futures. An organization may assess how climate policy, energy transition, demographic change and digital automation jointly affect supply chains, labour requirements and community relations. The aim is not to forecast one future with certainty, but to improve preparedness by testing strategic assumptions under uncertainty.

This approach resonates with complexity thinking. Organizations are embedded in systems where small changes may generate large effects and feedback loops are difficult to predict (Bateson, 1972; Mitchell, 2009). Scenario-based ESG intelligence acknowledges this complexity. It supports strategies that are robust across plausible futures rather than optimized for one expected path. It also helps identify opportunity spaces, such as circular innovation, inclusive services, sustainable urban partnerships or knowledge-intensive collaborations with communities.

Predictive ESG intelligence should therefore be integrated into strategic planning, risk management and governance reporting. Its value is highest when it changes decisions rather than merely producing dashboards. Boards and management teams can use predictive outputs to discuss vulnerabilities, allocate resources, redesign processes and engage stakeholders. At the same time, they must disclose the assumptions behind predictive models. Transparency is essential because future-oriented reporting can easily become speculative if uncertainty is understated.

Scenario work also has a behavioural dimension. It can reduce dependence on a single expected future, but it may reproduce cognitive biases when assumptions remain implicit (Ariely, 2008; Schoemaker, 1995). Predictive ESG intelligence should therefore include assumption registers, sensitivity analysis and structured disagreement. It should also consider the safe and just operating space in which economic activity occurs (Raworth, 2017). Its value emerges when sustainability supports innovation in products, processes and business models (Nidumolu et al., 2009; Schaltegger et al., 2017).

Table 1. From European reporting traditions to predictive ESG intelligence

Trajectory	Core knowledge logic	Contribution to ESG standards	AI-enabled extension
German	Measurement, standardisation and procedural discipline	Improves comparability, auditability and data reliability	Supports structured datasets and consistent predictive indicators
French	Public regulation, social responsibility and accountability	Links reporting to legitimacy, oversight and stakeholder expectations	Guides model use through transparency, responsibility and public purpose

Trajectory	Core knowledge logic	Contribution to ESG standards	AI-enabled extension
Italian	Systemic equilibrium and institutional continuity	Integrates economic, social and environmental dimensions of the firm	Connects ESG variables with resilience, continuity and scenario analysis

4.4. Illustrative example: climate transition as predictive ESG intelligence

The climate transition illustrates how a megatrend can become dynamic ESG insight and governance action. The initial focus is not a single environmental indicator, but a system of weak signals that may affect the organization over time: energy price volatility, regulatory pressure, carbon-related disclosure mandates, supplier vulnerability and changing consumer expectations.

In a descriptive framework, the organization may disclose current energy use or emissions. In a predictive ESG intelligence framework, the same information is linked to prospective exposure. A relevant indicator may include the proportion of strategic suppliers exposed to transition risk, with expected procurement or energy cost increases under different regulatory scenarios. The predictive output is not deterministic; it is a scenario-based estimate of cost increases, supply-chain vulnerabilities and implications for continuity.

The related governance decision may involve supplier diversification, energy-efficiency investment, revision of the transition strategy, contract renegotiation or stakeholder engagement. In summary: megatrend, climate transition; weak signals, energy price volatility, regulation and supply-chain exposure; dynamic metric, percentage of suppliers exposed to transition risk; predictive output, scenario-based estimate of cost increases; governance decision, supplier diversification and transition-plan revision. Predictive ESG intelligence is therefore a cognitive and organizational process that converts external transformations into measurable insights, scenarios and responsible action.

5 Discussion: Implications for Intelligent Organizations and Communities

The proposed framework has implications for both organizations and communities. For organizations, predictive ESG intelligence reframes sustainability reporting as a knowledge creation process. The organization does not simply collect information to satisfy external requirements. It learns how global transformations affect its business model, stakeholder relationships and capacity for continuity. This learning process can improve resilience because it enables earlier identification of vulnerabilities and more coherent responses to uncertainty (Zuboff, 1988).

For communities, the framework is relevant because organizational decisions shape social and urban ecosystems. Firms influence employment, skills, environmental quality, infrastructure use, innovation networks and quality of life. When organizations use megatrend analysis only for internal advantage, they may overlook broader effects. When they embed predictive intelligence in stakeholder dialogue, they can support shared value and inclusive growth. This matters in smart and hybrid environments, where work, public services and community interactions are increasingly mediated by digital technologies.

The framework also clarifies the relationship between human and artificial intelligence. AI can expand analytical capacity, but human intelligence remains responsible for interpretation, judgement and ethical direction. Collective intelligence is equally important because sustainability challenges cannot be solved within isolated hierarchies. Workers, experts, citizens, regulators and communities hold knowledge that may not be captured in formal datasets. Intelligent organizations must therefore design spaces for knowledge sharing, including hybrid workplaces, collaborative platforms and participatory governance mechanisms.

A further implication concerns compliance reductionism. ESG reporting can become formal when organizations focus only on producing documents. Predictive ESG intelligence resists this tendency by linking reporting with strategy and learning. However, it introduces risks. Models may create an illusion of certainty, automated systems may privilege measurable variables and organizations may overstate predictive capacity. These risks reinforce the need for governance oversight and for a clear distinction between analytical support and managerial responsibility.

From a practical perspective, organizations should begin with a diagnostic assessment of data, knowledge flows and governance routines. They should identify relevant megatrends, update indicators and establish interdisciplinary teams capable of interpreting predictive outputs. They should also document assumptions and expose them to challenge. This connects predictive ESG intelligence with corporate governance scholarship on board monitoring, strategic guidance and accountability (Zattoni and Pugliese, 2019). The objective is to make sustainability governance more informed, anticipatory and accountable (Kellogg, Valentine and Christin, 2020).

A final practical implication concerns algorithm and data governance. Predictive ESG systems should be assessed for explainability, fairness, auditability and proportionality, because technical opacity can weaken

accountability (Binns, 2018; Floridi et al., 2018). OECD principles link trustworthy AI with human-centred values, transparency, robustness and responsible stewardship (OECD, 2024). These requirements define when digital tools become legitimate organizational knowledge rather than unexamined technical authority (Kitchin, 2014; Orlikowski, 2007).

6 Conclusions

This paper has argued that the future of ESG standards depends on transforming megatrend research into predictive sustainability intelligence. Intelligent and sustainable organizations are not defined by digital adoption alone. They are defined by their ability to manage knowledge, interpret complexity and combine human and artificial intelligence in responsible value creation. The central contribution is the interpretation of ESG standards as dynamic knowledge artefacts and of AI as cognitive infrastructure for anticipatory sustainability governance.

The paper contributes to the IFKAD debate by connecting megatrends, sustainability standards, knowledge management and AI within a European intellectual framework. The analysis of German, French and Italian trajectories shows that contemporary sustainability reporting is grounded in complementary knowledge regimes: measurement, accountability and systemic equilibrium. These regimes converge in the need for reporting systems that are comparable, legitimate and capable of anticipating future conditions.

The proposed AI-enabled pathway moves through three phases: mapping megatrends and weak signals, embedding them into dynamic ESG metrics and developing scenario-based sustainability risk intelligence. The climate-transition example shows how this pathway translates a broad transformation into weak-signal interpretation, dynamic metrics, predictive scenarios and governance decisions. This supports the shift from descriptive disclosure to anticipatory governance and confirms that predictive systems strengthen resilience only when embedded in ethical oversight, stakeholder engagement and organizational learning.

The paper remains conceptual. It does not test the proposed framework empirically, nor does it compare specific AI tools or reporting systems. Future research could examine how organizations in different sectors implement predictive ESG analytics, how boards use forward-looking sustainability information and how communities participate in the interpretation of megatrends. Empirical studies could also explore whether predictive ESG intelligence improves resilience, innovation capacity and stakeholder dialogue.

7 Ethics Declaration

This paper is conceptual and does not report research involving human participants, personal data, confidential organizational records or experimental interventions. Ethical clearance was therefore not required.

8 AI Declaration

AI-assisted drafting support was used to structure and refine the paper from the authors' extended abstract and the conference template. The resulting text was reviewed and amended by the authors, who retain responsibility for the argument, citations, accuracy and final wording.

References

Ansoff, H.I. (1975) 'Managing strategic surprise by response to weak signals', California Management Review, Vol. 18, No. 2, pp. 21–33.

Antal, A.B. and Sobczak, A. (2007) 'Corporate social responsibility in France: A mix of national traditions and international influences', Business & Society, Vol. 46, No. 1, pp. 9–32.

Ariely, D. (2008) Predictably Irrational: The Hidden Forces That Shape Our Decisions, HarperCollins, New York.

Barney, J. (1991) 'Firm resources and sustained competitive advantage', Journal of Management, Vol. 17, No. 1, pp. 99–120.

Bateson, G. (1972) Steps to an Ecology of Mind, University of Chicago Press, Chicago.

Binns, R. (2018) 'Fairness in machine learning: Lessons from political philosophy', Proceedings of Machine Learning Research, Vol. 81, pp. 1–11.

Borgonovi, E. and Anessi Pessina, E. (2018) Economia delle aziende, Egea, Milan.

Cassandro, P.E. (1978) Impresa e responsabilità sociale, Il Mulino, Bologna.

Chauvey, J.N., Giordano-Spring, S., Cho, C.H. and Patten, D.M. (2015) 'The normativity and legitimacy of CSR disclosure: Evidence from France', Journal of Business Ethics, Vol. 130, No. 4, pp. 789–803.

Davenport, T.H. and Prusak, L. (1998) Working Knowledge: How Organizations Manage What They Know, Harvard Business School Press, Boston.

De Mauro, A., Greco, M. and Grimaldi, M. (2016) 'A formal definition of Big Data based on its essential features', Library Review, Vol. 65, No. 3, pp. 122–135.

EFRAG (2023) European Sustainability Reporting Standards: ESRS Set 1, EFRAG, Brussels.

Elkington, J. (1997) Cannibals with Forks: The Triple Bottom Line of 21st Century Business, Capstone, Oxford.
European Commission (2022) Corporate Sustainability Reporting Directive (CSRD), European Commission, Brussels.
European Commission (2023) Commission Delegated Regulation (EU) 2023/2772 on European Sustainability Reporting Standards, Official Journal of the European Union, Brussels.
Ferrero, G. (1968) La teoria economica dell'azienda, Giuffrè, Milan.
Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V. et al. (2018) 'AI4People: An ethical framework for a good AI society', Minds and Machines, Vol. 28, No. 4, pp. 689–707.
Freeman, R.E. (1984) Strategic Management: A Stakeholder Approach, Pitman, Boston.
Geels, F.W. (2002) 'Technological transitions as evolutionary reconfiguration processes', Research Policy, Vol. 31, Nos. 8–9, pp. 1257–1274.
George, G., Howard-Grenville, J., Joshi, A. and Tihanyi, L. (2016) 'Understanding and tackling societal grand challenges through management research', Academy of Management Journal, Vol. 59, No. 6, pp. 1880–1895.
German Council for Sustainable Development (2020) The Sustainability Code: Benchmarking Sustainable Business, German Council for Sustainable Development, Berlin.
Grant, R.M. (1996) 'Toward a knowledge-based theory of the firm', Strategic Management Journal, Vol. 17, Winter Special Issue, pp. 109–122.
GRI (2021) Consolidated Set of the GRI Standards, Global Reporting Initiative, Amsterdam.
Hahn, T., Pinkse, J., Preuss, L. and Figge, F. (2010) 'Tensions in corporate sustainability: Towards an integrative framework', Journal of Business Ethics, Vol. 99, No. 2, pp. 217–235.
Hastie, T., Tibshirani, R. and Friedman, J. (2009) The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Springer, New York.
Hoffmann, E., Dietsche, C. and Hobelsberger, C. (2018) 'Between mandatory and voluntary: non-financial reporting by German companies', Sustainability Management Forum, Vol. 26, No. 1, pp. 47–63.
IFRS Foundation (2023a) IFRS S1 General Requirements for Disclosure of Sustainability-related Financial Information, IFRS Foundation, London.
IFRS Foundation (2023b) IFRS S2 Climate-related Disclosures, IFRS Foundation, London.
ISO (2010) ISO 26000: Guidance on Social Responsibility, International Organization for Standardization, Geneva.
Kellogg, K.C., Valentine, M.A. and Christin, A. (2020) 'Algorithms at work: The new contested terrain of control', Academy of Management Annals, Vol. 14, No. 1, pp. 366–410.
Kitchin, R. (2014) The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences, Sage, London.
Mitchell, M. (2009) Complexity: A Guided Tour, Oxford University Press, Oxford.
Nidumolu, R., Prahalad, C.K. and Rangaswami, M.R. (2009) 'Why sustainability is now the key driver of innovation', Harvard Business Review, Vol. 87, No. 9, pp. 56–64.
Nonaka, I. and Takeuchi, H. (1995) The Knowledge-Creating Company, Oxford University Press, Oxford.
OECD (2024) Recommendation of the Council on Artificial Intelligence, OECD/LEGAL/0449, as amended in 2024, OECD Legal Instruments, Paris.
Onida, P. (1971) Economia d'azienda, UTET, Turin.
Orlikowski, W.J. (2007) 'Sociomaterial practices: Exploring technology at work', Organization Studies, Vol. 28, No. 9, pp. 1435–1448.
Porter, M.E. and Kramer, M.R. (2011) 'Creating shared value', Harvard Business Review, Vol. 89, Nos. 1–2, pp. 62–77.
Raworth, K. (2017) Doughnut Economics: Seven Ways to Think Like a 21st-Century Economist, Random House Business, London.
Rohrbeck, R. and Kum, M.E. (2018) 'Corporate foresight and its impact on firm performance: A longitudinal analysis', Technological Forecasting and Social Change, Vol. 129, pp. 105–116.
Russell, S. and Norvig, P. (2021) Artificial Intelligence: A Modern Approach, 4th ed., Pearson, Harlow.
Schaltegger, S., Hansen, E.G. and Lüdeke-Freund, F. (2017) 'Business models for sustainability', Organization & Environment, Vol. 29, No. 1, pp. 3–10.
Schoemaker, P.J.H. (1995) 'Scenario planning: A tool for strategic thinking', Sloan Management Review, Vol. 36, No. 2, pp. 25–40.
Task Force on Climate-related Financial Disclosures (2017) Recommendations of the Task Force on Climate-related Financial Disclosures, Financial Stability Board, Basel.
Teece, D.J., Pisano, G. and Shuen, A. (1997) 'Dynamic capabilities and strategic management', Strategic Management Journal, Vol. 18, No. 7, pp. 509–533.
Van der Byl, C.A. and Slawinski, N. (2015) 'Embracing tensions in corporate sustainability', Organization & Environment, Vol. 28, No. 1, pp. 54–79.
Vial, G. (2019) 'Understanding digital transformation: A review and a research agenda', Journal of Strategic Information Systems, Vol. 28, No. 2, pp. 118–144.
Weick, K.E. (1995) Sensemaking in Organizations, Sage.
World Commission on Environment and Development (1987) Our Common Future, Oxford University Press, Oxford.
Zappa, G. (1927) Il reddito d'impresa, Giuffrè, Milan.
Zattoni, A. and Pugliese, A. (2019) Corporate Governance, Egea, Milan.
Zuboff, S. (1988) In the Age of the Smart Machine: The Future of Work and Power, Basic Books.

The Experience of Error in Business Administration: The Case of Predictive Assessment in the Early Emergence of Corporate Distress

Domenico Celenza[1], Alessia Nardecchia[2], Francesco Di Vona[2], Marco Sebastiano[2]

[1] University of Cassino and Southern Lazio, Cassino, Italy

[2] Independent scholars

Abstract

This paper examines the experience of error in business-administrative judgement as a constitutive dimension of knowledge creation and as a critical object of governance in intelligent and sustainable organisations. The central thesis is that error is not a residual disturbance to be eliminated by predictive technologies, but a structural component of organisational knowledge that must be governed through hybrid human-AI arrangements. The analysis focuses on predictive assessment in the early emergence of corporate distress, a field in which managerial judgement, financial information, organisational arrangements and algorithmic models interact under conditions of uncertainty.

The paper argues that organisational intelligence cannot be reduced to the adoption of digital technologies or artificial intelligence systems. It depends on the capacity of organisations to recognise fallibility, interpret weak signals, transform mistakes into shared knowledge and design decision routines that strengthen resilience over time. The duty to establish adequate organisational, administrative and accounting arrangements under the Italian Code of Corporate Crisis and Insolvency is interpreted as an institutional response to the structural limits of judgement. Artificial intelligence can reinforce this infrastructure by processing large volumes of heterogeneous data, detecting non-obvious patterns and supporting early-warning systems. At the same time, predictive technologies introduce new forms of error, including poor data quality, model opacity, automation bias and the displacement of responsibility from human decision-makers to technical systems.

The paper therefore proposes a detection-interpretation-decision-learning framework of hybrid predictive governance. In this framework, AI supports the identification of patterns, weak signals and anomalies; human judgement interprets the organisational and strategic context; boards and management assume responsibility for decisions; model and organisational errors are reviewed; and the system learns and improves over time. The contribution is conceptual. It advances the debate on intelligent organisations and communities by showing that technology does not make organisations intelligent by itself. Organisations become intelligent when they develop the cultural, informational and technological capacity to learn from error before it becomes irreversible corporate failure.

Keywords: Intelligent organisations; knowledge management; artificial intelligence; corporate distress; predictive governance; organisational learning.

Paper type: Academic Research Paper.

1 Introduction

Sustainable value creation increasingly depends on organisations that can combine knowledge, technology and responsibility. Digital transformation, artificial intelligence and data-intensive control systems are often expected to make organisations more responsive, resilient and capable of anticipating discontinuities. Yet this expectation remains incomplete if it does not address a basic condition of managerial action: judgement is fallible. Firms do not operate in transparent environments. They interpret incomplete signals, rely on provisional models, select certain indicators while neglecting others, and construct narratives of continuity that may later prove fragile. An intelligent organisation is therefore not an organisation without error. It is an organisation that can recognise, process and govern error as part of its learning capacity.

This paper examines this issue through the specific case of predictive assessment in the early emergence of corporate distress. Corporate distress rarely appears as a sudden and isolated event. It typically develops through weak signals, such as declining margins, liquidity tensions, delayed payments, deteriorating customer relationships, operational rigidity, worsening credit conditions or the progressive loss of strategic coherence. These signals may be observable before insolvency becomes evident, but availability is not equivalent to knowledge. Data become useful only when they are interpreted, connected to organisational memory, discussed in appropriate decision arenas and translated into timely action. The problem is therefore not merely informational. It is cognitive, organisational and technological at the same time.

The central thesis of the paper is that error is not a residual disturbance to be eliminated by predictive technologies, but a structural component of organisational knowledge that must be governed through hybrid

human-AI arrangements. This thesis shifts attention from the ambition to remove error through automation to the design of governance systems capable of making error visible, interpretable and learnable. Predictive technologies may improve detection, but they cannot decide by themselves what a signal means, which trade-offs are acceptable or which response is responsible. The organisational challenge is not to delegate judgement to AI, but to build arrangements in which artificial intelligence, human judgement and collective governance correct and strengthen one another (Raisch and Krakowski, 2021).

In such hybrid arrangements, AI helps detect patterns, weak signals and anomalies; human judgement interprets the organisational and strategic context; boards and management assume responsibility for decisions; model and organisational errors are reviewed; and the system learns and improves over time. The paper translates this logic into a detection-interpretation-decision-learning framework, which is introduced at the beginning of the manuscript and then developed throughout the analysis. The framework is intended as an interpretive map rather than as a rigid procedural checklist. Its purpose is to show how predictive governance can transform error from a source of failure into a source of organisational learning.

The argument develops from a philosophical and epistemological understanding of error and reinterprets it in the context of artificial intelligence and predictive governance. Rather than treating error as an external disturbance to be removed by increasingly sophisticated technologies, the paper considers fallibility as a constitutive dimension of knowledge creation, organisational learning and managerial judgement. Selected epistemological accounts show that error is not external to knowledge, but is part of the process through which assumptions are tested, corrected and revised. This perspective is then connected to cognitive psychology, which demonstrates that managerial judgement is systematically exposed to biases such as anchoring, availability, representativeness and overconfidence (Tversky and Kahneman, 1974; Kahneman, 2011).

The Italian tradition of Economia Aziendale further reinforces the argument by interpreting the firm as a systemic and unitary phenomenon, in which economic, financial, patrimonial and organisational dimensions must be evaluated jointly (Zappa, 1927; Onida, 1971; Ferrero, 1968; Amaduzzi, 1963). Finally, studies on knowledge management and artificial intelligence explain how organisations can transform dispersed information into collective intelligence, while also showing that digital systems may introduce new vulnerabilities, including data bias, model opacity, automation bias and the displacement of responsibility from human decision-makers to algorithmic infrastructures (Nonaka and Takeuchi, 1995; Grant, 1996; Davenport and Prusak, 1998; Russell and Norvig, 2021).

2 A Framework of Hybrid Predictive Governance

The proposed framework places the detection-interpretation-decision-learning cycle at the core of predictive governance. Its first assumption is that corporate distress is not only a financial condition, but also a knowledge failure. Relevant information may be absent, delayed, fragmented, misinterpreted or politically suppressed. Predictive systems can increase the visibility of weak signals; however, the meaning of those signals remains context-dependent. For this reason, organisations must integrate algorithmic detection with human interpretation, responsible decision-making and systematic learning from predictive failures.

The framework consists of four interdependent layers. The first layer is detection: which weak signals are visible, and which remain obscured? The second layer is interpretation: what do those signals mean within the specific organisational, strategic and financial context of the firm? The third layer is decision: who is required to act, when, and with what level of responsibility? The fourth layer is learning: what did the prediction get right or wrong, and how should the organisation revise its data, models, assumptions and routines? The framework therefore rejects both technological determinism and purely intuitive judgement. It conceives predictive governance as a hybrid collaboration between human and artificial intelligence.

Figure 1 summarises the logic of the framework. AI expands the organisation's capacity to detect patterns, weak signals and anomalies. Human judgement connects those signals with tacit knowledge, strategic context and ethical considerations. Boards and management remain responsible for decisions because algorithmic outputs are inputs into governance, not substitutes for governance. Finally, errors are reviewed not only to allocate responsibility, but also to improve the system. False positives, false negatives, ignored warnings and delayed reactions become learning events.

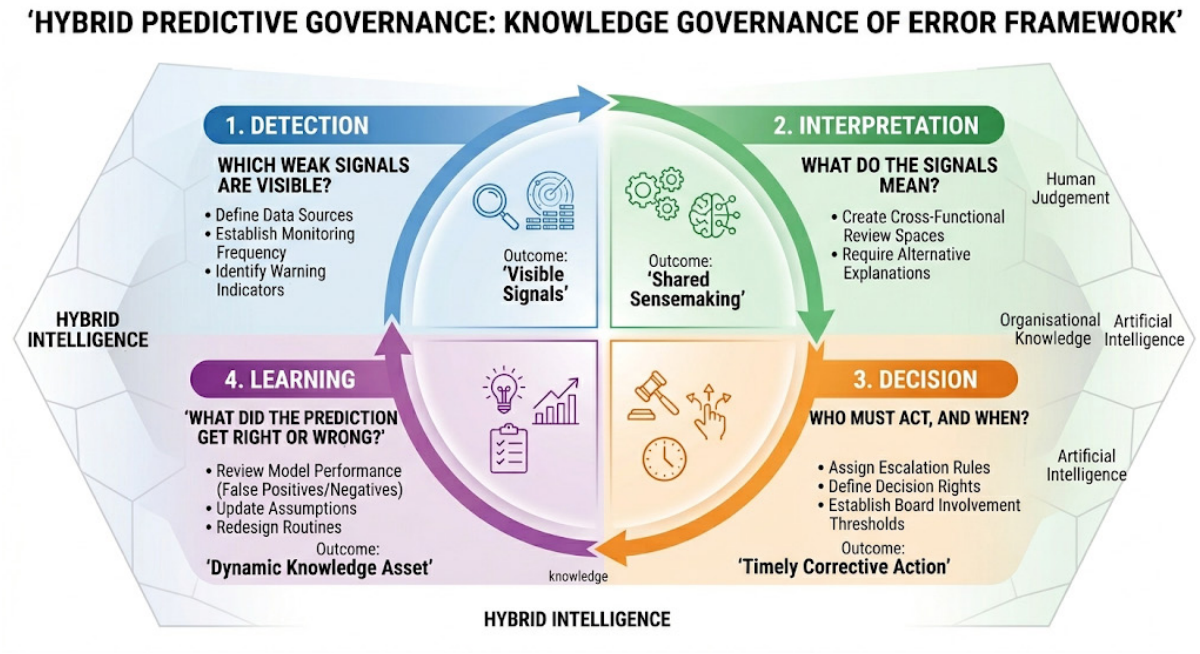


Figure 1. Hybrid predictive governance: knowledge governance of error framework.

Table 1. Knowledge governance of error in predictive assessment

Layer	Main question	Typical error	Governance response
Detection	Which weak signals are visible?	Relevant data are missing, delayed or fragmented.	Define data sources, monitoring frequency and warning indicators.
Interpretation	What do the signals mean?	Signals are normalised, rationalised or read through biased narratives.	Create cross-functional review spaces and require alternative explanations.
Decision	Who must act, and when?	Knowledge does not translate into timely corrective action.	Assign escalation rules, decision rights and board-involvement thresholds.
Learning	What did the prediction get right or wrong?	False positives and false negatives remain unexplained.	Review model performance, update assumptions and redesign routines.

Table 1 presents the same framework in analytical form. It highlights the main question, typical error and governance response associated with each layer.

3 Error as an Organisational Knowledge Experience

The experience of error is relevant for organisational theory because every managerial decision presupposes a representation of reality, and every representation may be partial. Organisations decide under uncertainty. They simplify complexity, select indicators, rely on forecasts and interpret signals through existing narratives. Error therefore cannot be understood only as an individual mistake. It often reveals the limits of the organisational knowledge system: what the organisation observes, what it ignores, how it interprets information and how it reacts when evidence contradicts expectations.

A concise epistemological reading is useful here insofar as it supports the organisational argument. Popper's view of knowledge as conjectural and exposed to falsification is particularly relevant (Popper, 1972). Knowledge grows when assumptions are tested and revised. In organisations, the same principle applies when forecasts, dashboards, risk classifications or managerial narratives are compared with subsequent events. A wrong prediction may be damaging, but it can also reveal why organisational assumptions were incomplete. Error becomes productive when it is not hidden, normalised or reduced to blame, but used to revise the system that generated it.

In organisational terms, error becomes meaningful when it is investigated as a symptom of a wider knowledge process. A wrong forecast may reveal poor data quality. A delayed reaction may reveal hierarchical silence. A distorted assessment may reveal incentives to preserve optimistic narratives. An ignored warning may reveal the absence of legitimate spaces for contradiction. The epistemology of error therefore becomes a theory of organisational learning. The organisation learns when it can move from the occurrence of error to its explanation, and from explanation to redesign.

This reasoning is close to the distinction proposed by Argyris and Schon (1978) between single-loop and double-loop learning. Single-loop learning corrects action while preserving the assumptions that generated it. Double-loop learning questions those assumptions. In the field of corporate distress, single-loop learning may add a new financial indicator or increase reporting frequency. Double-loop learning asks why existing indicators were not taken seriously, why weak signals were normalised and why decision-makers delayed corrective intervention. It is this second level that turns error into a strategic knowledge asset.

The perspective also resonates with knowledge-management theory. Nonaka and Takeuchi (1995) emphasise that knowledge creation depends on the interaction between tacit and explicit knowledge. Grant (1996) interprets the firm as an institution for integrating specialised knowledge. Davenport and Prusak (1998) similarly stress that organisational knowledge requires practices, processes and contexts that allow information to become actionable. These approaches suggest that organisational intelligence does not reside simply in databases or analytical tools. It lies in the capacity to connect experience, interpretation and action. Error is one of the occasions through which this connection can be tested. If an organisation suppresses or hides errors, it impoverishes its knowledge base. If it analyses them, it increases its capacity to perceive emerging vulnerabilities.

4 Managerial Judgement, Bias and Corporate Distress

Corporate distress is a useful field for observing the organisational consequences of judgemental error. The prediction of failure has a long tradition in accounting and finance. Beaver (1966, 1968) and Altman (1968) showed that financial ratios can provide meaningful indications of impending failure. Later models refined the analysis through hazard models, international comparisons and machine-learning approaches (Shumway, 2001; Altman et al., 2017). These contributions are essential, but they do not remove the role of judgement. Indicators must be selected, interpreted, contextualised and acted upon. The presence of a warning signal does not guarantee that the organisation will understand it.

Cognitive psychology explains why this gap persists. Tversky and Kahneman (1974) demonstrated that judgement under uncertainty is shaped by heuristics. Heuristics are useful because they reduce complexity, but they may also generate systematic biases. Anchoring can lead managers to interpret current performance through the lens of past success. Availability can make recent or emotionally salient events more influential than structural trends. Representativeness can encourage decision-makers to treat a temporary recovery as evidence that the business model remains robust. Overconfidence can lead management to underestimate both the probability and the speed of deterioration.

These biases are particularly dangerous in the early stages of crisis. A firm may still be operating, selling, paying wages and maintaining relationships with banks and suppliers. The visible continuity of operations can conceal a progressive erosion of economic equilibrium. Managers may treat volatility as noise, temporary liquidity shortages as manageable timing problems or commercial decline as a reversible market fluctuation. The more the organisation has invested its identity in a narrative of competence and control, the more difficult it becomes to accept that weak signals may be symptoms of deeper imbalance.

The psychology of error must therefore be connected to the sociology of organisations. Biases do not operate only inside individual minds. They are reinforced or weakened by organisational routines, incentive systems and power relations. A board that receives highly aggregated reports may be formally informed but substantively unable to understand the sources of risk. A chief financial officer may identify anomalies but fail to obtain attention if the dominant strategic narrative is expansionary. Middle managers may possess tacit knowledge of operational fragility, but that knowledge may not reach the governance level. Error is thus not only cognitive. It is distributed across information flows, reporting formats and decision rights.

Simon (1997) described administrative behaviour as boundedly rational. Decision-makers search, simplify and satisfice because perfect rationality is impossible. The issue is not to replace bounded rationality with an ideal of total optimisation, but to design arrangements that compensate for its limits. In the context of distress prediction, this means constructing processes that make weak signals visible, require competing interpretations and reduce the dependence of strategic decisions on intuition alone. The aim is not to distrust managerial experience, but to embed that experience in a corrective architecture of knowledge.

For this reason, the language of error should not be reduced to blame. Blame may be appropriate in cases of negligence, opportunism or concealment, but a generalised culture of blame can weaken learning by encouraging defensive behaviour. The relevant question for intelligent organisations is not only who made a mistake, but also what conditions made the mistake plausible, persistent or invisible. This shift is essential for organisations that aspire to be sustainable. Sustainability requires continuity of value creation, and continuity requires the ability to detect deterioration before it becomes irreversible.

5 “Economia Aziendale” and Adequate Arrangements

The Italian tradition of “Economia Aziendale” offers an important theoretical basis for linking error, knowledge and governance. Zappa (1927, 1956) placed the firm within a unitary and dynamic conception of economic activity. Onida (1971), Amaduzzi (1963) and Ferrero (1968) developed the idea that business administration must consider the firm as a system in which economic, financial, patrimonial and organisational aspects are interdependent. More recent contributions have confirmed the relevance of this integrated view for understanding the firm as a complex institution of value creation and accountability (Borgonovi and Anessi Pessina, 2018). The assessment of corporate conditions cannot be reduced to isolated ratios or single episodes. It requires an integrated judgement that connects income generation, financial equilibrium, assets, liabilities, organisational structures and strategic direction.

From this perspective, error often emerges from partiality. The evaluator may focus on profitability while neglecting cash generation, observe liquidity while neglecting debt sustainability, read accounting data without understanding operational processes, or treat organisational disorder as secondary to financial measurement. Such partiality is especially problematic in crisis prevention because distress is systemic. It develops when multiple dimensions of the firm cease to support one another. A knowledge-driven assessment must therefore preserve the unity of the business phenomenon while still using specialised indicators.

The Italian Code of Corporate Crisis and Insolvency gives this problem a specific institutional form. The duty to establish adequate organisational, administrative and accounting arrangements can be interpreted as a legal and managerial response to fallibility. Adequacy does not merely require the existence of formal procedures. It requires the capacity to detect imbalances, monitor business continuity, produce reliable information and support timely intervention. In this sense, the law translates an epistemological problem into a governance requirement: because judgement is fallible, organisations must create arrangements that reduce the probability that relevant signals remain unseen or misunderstood.

The notion of adequacy is therefore close to knowledge management. Adequate arrangements are not only controls in a narrow compliance sense. They are infrastructures for producing, validating and circulating knowledge about the firm. They define what information is collected, how it is processed, who receives it, which thresholds trigger attention, how alternative explanations are discussed and how responsibility for action is allocated. Their quality should be assessed not only by their formal existence, but by their capacity to support intelligent interpretation.

This reading is also consistent with contemporary debates on corporate governance. Boards are increasingly expected to oversee not only financial performance but also risk, sustainability, digital transformation and organisational resilience (Zattoni and Pugliese, 2019). However, board oversight is effective only if the board receives information that is timely, intelligible and connected to the strategic context. A dashboard that produces numerous indicators without meaningful interpretation may create a false impression of control. Conversely, a well-designed governance process can use indicators as starting points for informed questioning.

Adequacy thus has a relational dimension. It depends on the interaction among managers, accountants, controllers, auditors, directors and, where relevant, external advisers. Each actor holds partial knowledge. The task of governance is to connect these fragments into a sufficiently coherent view of the firm's condition. The early emergence of distress is precisely the situation in which this connection matters most. At that stage, the organisation still has choices. The quality of knowledge integration can determine whether those choices are recognised in time.

6 Artificial Intelligence in Early-Warning Systems

Artificial intelligence changes the relationship between error and judgement. Predictive algorithms can process volumes of data that exceed the capacity of individual decision-makers. They can detect non-linear relationships, compare current patterns with historical cases, update predictions as new information arrives and generate early warnings that may not be visible through traditional reporting. In the field of distress prediction, machine-learning tools can integrate financial ratios, payment behaviour, market data, textual information, operational indicators and external signals. Used appropriately, they expand the perceptual capacity of the organisation.

This expansion is important for intelligent and sustainable organisations. Sustainability is not only environmental or social; it also concerns the durability of value creation and the capacity to preserve productive relationships with employees, creditors, suppliers, customers and communities. Early-warning systems contribute to this objective by increasing the probability of timely intervention. They can support turnaround actions, debt restructuring, strategic repositioning or the redesign of business processes before value destruction becomes irreversible (Guatri, 1995).

However, artificial intelligence does not eliminate error. It relocates and transforms it. Data-quality problems may lead models to learn from incomplete, obsolete or biased information. Model risk may arise when an algorithm performs well on historical data but poorly in new circumstances. Opacity may prevent users from understanding why a warning has been produced. Automation bias may induce decision-makers to accept the algorithmic output without sufficient scrutiny. Conversely, algorithm aversion may lead experienced managers to reject useful warnings because they conflict with intuition or established narratives. The intelligent organisation must govern both forms of bias: excessive trust in the model and defensive rejection of the model.

The problem is not merely technical. It is organisational and ethical. Predictive systems influence attention and may affect access to credit, restructuring decisions and relationships among stakeholders. If the model is treated as neutral, responsibility may be displaced from human judgement to technical infrastructure. Yet algorithms are designed, trained, selected and interpreted by organisations. Their outputs are not decisions; they are inputs into decision processes. For this reason, AI governance must be integrated with corporate governance and knowledge governance.

Russell and Norvig (2021) define artificial intelligence in terms of systems that perceive, reason and act. In organisations, however, perception, reasoning and action are distributed across people and technologies. An early-warning algorithm may perceive patterns, but human actors must decide whether those patterns are strategically meaningful. It may generate a probability of distress, but it cannot by itself determine what level of risk is acceptable, which trade-offs are legitimate or how the consequences of intervention should be allocated. These are managerial and ethical questions.

The central issue is therefore the design of human-in-the-loop arrangements. Such arrangements should specify the data used by the model, the frequency of monitoring, the thresholds for escalation, the criteria for explanation, the responsibility for review and the procedures for challenging or validating the output. They should also include feedback loops. If a model produces false positives or false negatives, the organisation should not simply record the event. It must investigate why the error occurred and whether it reflects data limitations, model design, contextual change or interpretive failure. In this way, algorithmic error becomes part of organisational learning rather than a hidden technical defect.

The solution is therefore not to let AI decide everything (Parasuraman and Manzey, 2010). The solution is to build hybrid arrangements in which AI identifies signals that humans may miss, while human and collective judgement interpret those signals responsibly. This hybrid logic is particularly important in the early emergence of corporate distress, where signals are ambiguous, consequences are distributed among multiple stakeholders and decisions must often be taken before certainty is available.

7 Discussion: Intelligent, Sustainable and Inclusive Organisations

The proposed framework contributes to the debate on intelligent and sustainable organisations by shifting attention from technological adoption to knowledge governance. Organisations often describe themselves as intelligent because they use advanced analytics, dashboards or AI-based platforms. Such technologies are useful, but they do not by themselves create intelligence. Intelligence is relational. It emerges when information is made available to the right actors, when interpretation can be challenged, when decisions are connected to evidence and when the outcomes of decisions are used for learning.

This point has consequences for sustainability. Sustainable organisations must manage tensions among economic performance, social responsibility and long-term resilience (Hahn et al., 2010; Van der Byl and Slawinski, 2015). Corporate distress is not only a private problem of shareholders. It can affect employment, suppliers, creditors, local communities and the continuity of productive knowledge. Early-warning systems, if properly governed, can therefore support a broader form of value preservation. They allow organisations to intervene before crisis destroys relationships and before restructuring options become narrow.

The argument also has implications for inclusive innovation. A predictive governance system should not centralise intelligence exclusively in top management or technical experts. It should recognise that relevant knowledge is distributed. Employees may observe operational inefficiencies before they appear in accounts. Sales teams may perceive customer fragility before it affects revenues. Accountants may identify payment irregularities before management accepts their strategic significance. External stakeholders may provide signals that internal actors ignore. An inclusive intelligent organisation creates channels through which such knowledge can enter the assessment process.

Digital and hybrid work environments make this objective both easier and more difficult. They can facilitate the rapid sharing of information across geographical boundaries and organisational units. They can also fragment attention, multiply reports and weaken informal interpretation. The challenge is to build digital meeting spaces that support collective sensemaking. A dashboard is not a community of interpretation. A community of interpretation requires trust, accountability and the possibility of raising inconvenient questions without immediate defensive reaction.

The framework also suggests that AI ethics should be connected to organisational learning. Many discussions of AI ethics focus on fairness, transparency and accountability. These principles are essential, but they must be translated into everyday governance practices. In early-warning systems, transparency means that decision-makers understand the main drivers of a prediction. Accountability means that the organisation can identify who reviewed the output and what action was taken. Fairness means that the consequences of predictive classifications are considered in relation to stakeholders. Learning means that errors in prediction are not hidden, but used to improve the system.

Finally, the paper clarifies a paradox of intelligent organisations. The more an organisation aspires to be intelligent, the more it must institutionalise doubt. Confidence in technology, strategy or past success can become a source of blindness. Intelligent organisations are therefore not those that declare certainty, but those that maintain the capacity to revise themselves. They design routines that make error discussable, data contestable and judgement accountable. This is not a sign of weakness. It is a condition of resilience.

7 Conclusions

This paper has developed a conceptual account of the experience of error in business-administrative judgement, with particular reference to predictive assessment in the early emergence of corporate distress. Its central thesis is that error is not a residual disturbance to be removed by automation, but a structural dimension of organisational knowledge that must be governed. The paper has argued that predictive technologies can improve the detection of weak signals, but they cannot eliminate the need for human interpretation, board responsibility and organisational learning.

The contribution of the paper is threefold. First, it reframes error as a knowledge asset for intelligent organisations. Error indicates where information, interpretation and action have failed to connect. Second, it interprets adequate organisational, administrative and accounting arrangements as knowledge infrastructures aimed at governing managerial fallibility. Third, it proposes a detection-interpretation-decision-learning framework of hybrid predictive governance in which artificial intelligence supports, but does not replace, human and organisational judgement.

The practical implication is that organisations should not evaluate early-warning systems only by their technical accuracy. They should also assess whether those systems improve the circulation of knowledge, clarify responsibility, support timely decision-making and generate learning from false positives and false negatives. A predictive model that is accurate but ignored is not useful. A model that is trusted blindly is dangerous. The organisational challenge is to create a balanced relationship between computational intelligence and critical judgement.

The paper is limited by its conceptual nature. It does not test the proposed framework empirically. Future research could examine how firms design adequate arrangements, how boards interpret AI-generated warnings, how different organisational cultures respond to prediction errors and how stakeholders are affected by early-warning classifications. Comparative studies across legal systems and industries would also be valuable. Nonetheless, the theoretical argument is clear: intelligent and sustainable organisations are not defined by the absence of error. They are defined by the capacity to learn from error before it becomes organisational decline.

8 Ethics Declaration

This paper is conceptual and does not report research involving human participants, personal data, experiments, interviews or confidential organisational records. Ethical clearance was therefore not required. Should the framework be applied in future empirical research, the authors will seek the necessary ethical approval and ensure compliance with applicable rules on confidentiality, informed consent and data protection.

9 AI Declaration

An AI writing assistant was used as a drafting and language-support tool to expand the extended abstract into a full paper, improve coherence and check alignment with the conference template. The conceptual framing, selection of sources, legal interpretation, final judgement and responsibility for the submitted text remain with the authors. AI-generated material should be reviewed, edited and verified before submission.

References

Altman, E.I. (1968) "Financial ratios, discriminant analysis and the prediction of corporate bankruptcy", The Journal of Finance, Vol. 23, No. 4, pp. 589-609.

Altman, E.I., Iwanicz-Drozdowska, M., Laitinen, E.K. and Suvas, A. (2017) "Financial distress prediction in an international context: A review and empirical analysis of Altman's Z-score model", Journal of International Financial Management & Accounting, Vol. 28, No. 2, pp. 131-171.
Amaduzzi, A. (1963) L'azienda nel suo sistema e nell'ordine delle sue rilevazioni, Giuffre, Milano.
Argyris, C. and Schon, D.A. (1978) Organizational Learning: A Theory of Action Perspective, Addison-Wesley, Reading, MA.
Beaver, W.H. (1966) "Financial ratios as predictors of failure", Journal of Accounting Research, Vol. 4, pp. 71-111.
Beaver, W.H. (1968) "Market prices, financial ratios, and the prediction of failure", Journal of Accounting Research, Vol. 6, No. 2, pp. 179-192.
Borgonovi, E. and Anessi Pessina, E. (2018) Economia delle aziende, Egea, Milano.
Davenport, T.H. and Prusak, L. (1998) Working Knowledge: How Organizations Manage What They Know, Harvard Business School Press, Boston.
Decreto legislativo 12 gennaio 2019, n. 14, Codice della crisi d'impresa e dell'insolvenza, Gazzetta Ufficiale della Repubblica Italiana.
Ferrero, G. (1968) La teoria economica dell'azienda, Giuffre, Milano.
Grant, R.M. (1996) "Toward a knowledge-based theory of the firm", Strategic Management Journal, Vol. 17, Winter Special Issue, pp. 109-122.
Guatri, L. (1995) Turnaround: declino, crisi e ritorno al valore, Egea, Milano.
Hahn, T., Pinkse, J., Preuss, L. and Figge, F. (2010) "Tensions in corporate sustainability: Towards an integrative framework", Journal of Business Ethics, Vol. 99, No. 2, pp. 217-235.
Kahneman, D. (2011) Thinking, Fast and Slow, Farrar, Straus and Giroux, New York.
Nonaka, I. and Takeuchi, H. (1995) The Knowledge-Creating Company, Oxford University Press, Oxford.
Onida, P. (1971) Economia d'azienda, UTET, Torino.
Parasuraman, R. and Manzey, D. (2010) "Complacency and automation bias in the human use of automation: a review", Human Factors, Vol. 52, No. 3, pp. 381-410.
Popper, K. (1972) Objective Knowledge: An Evolutionary Approach, Oxford University Press, Oxford.
Raisch, S. and Krakowski, S. (2021) "Artificial intelligence and management: The automation-augmentation paradox", Academy of Management Review, Vol. 46, No. 1, pp. 192-210.
Russell, S. and Norvig, P. (2021) Artificial Intelligence: A Modern Approach, 4th ed., Pearson, Harlow.
Shumway, T. (2001) "Forecasting bankruptcy more accurately: A simple hazard model", The Journal of Business, Vol. 74, No. 1, pp. 101-124.
Simon, H.A. (1997) Administrative Behavior, 4th ed., Free Press, New York.
Tversky, A. and Kahneman, D. (1974) "Judgment under uncertainty: Heuristics and biases", Science, Vol. 185, No. 4157, pp. 1124-1131.
Van der Byl, C.A. and Slawinski, N. (2015) "Embracing tensions in corporate sustainability", Organization & Environment, Vol. 28, No. 1, pp. 54-79.
Zappa, G. (1927) Il reddito d'impresa, Giuffre, Milano.
Zappa, G. (1956) Le produzioni nell'economia delle imprese, Giuffre, Milano.
Zattoni, A. and Pugliese, A. (2019) Corporate Governance, Egea, Milano.

Research on AI and Sustainability Ratings: Evidence from the Noverim Case

Domenico Celenza[1], Marco Sebastiano[2], Andrea Maturo[2], Antonella Fortuna[1]
[1] University of Cassino and Southern Lazio, Cassino, Italy
[2] Independent scholar

Abstract

The convergence between artificial intelligence, sustainability measurement and knowledge management is reshaping how organisations diagnose vulnerability, create value and build resilience. This paper examines the Noverim case as the basis for a study hypothesis: AI-enabled sustainability ratings may help identify organisational disequilibrium before conventional accounting data reveal it. The article does not present original empirical evidence. It develops a conceptual and case-based framework centred on an intelligent predictive system for assessing sustainability and business equilibrium. The reference to 500 Italian knowledge-intensive firms is used solely as a prospective field for future validation; it is not an observed sample, a tested dataset or evidence supporting the hypothesis. The study asks how heterogeneous knowledge assets, sustainability ratings and predictive analytics can be integrated to support early diagnosis, resilience-building and knowledge-driven innovation. The proposed framework interprets sustainability as an endogenous component of equilibrium, since ESG and organisational conditions affect risk exposure, adaptive capacity, stakeholder trust and the durability of value creation. The Noverim architecture combines supervised machine-learning techniques with unsupervised behavioural clustering to identify latent risk profiles and possible trajectories of imbalance. On this basis, the paper formulates three propositions: first, such systems may anticipate disequilibrium over a 12-to-18-month horizon; second, weak signals such as reduced intangible investment, turnover of key personnel, deteriorating internal sustainability indicators and strategy-conduct misalignment may reveal early fragility; third, intelligent ratings may help distinguish reversible from structural crises. These propositions require future empirical testing. The paper contributes to the literature on knowledge-based organisations by conceptualising predictive intelligence as a knowledge infrastructure for organisational learning, resilience mapping and accountable decision-making.

Keywords: predictive analytics, sustainability ratings, corporate equilibrium, organisational resilience, machine learning

Paper type: Academic Research Paper

1 Introduction

Intelligent organisations are increasingly expected to combine economic performance, social responsibility and the capacity to learn from data. This expectation is particularly relevant for knowledge-intensive firms, whose competitiveness depends on specialised expertise, organisational routines, relational capital, digital capabilities and the interpretation of weak signals. In these firms, distress rarely begins as a purely accounting event. It usually appears through dispersed and initially ambiguous symptoms: a slowdown in knowledge investment, the loss of key personnel, inconsistent strategic choices, fragile cash-flow planning or a declining capacity to translate sustainability commitments into managerial practice.

This issue is directly connected with the debate on intelligent and sustainable organisations. Digital transformation has expanded the volume, velocity and variety of data available to firms, but data availability does not automatically generate organisational intelligence. Intelligence emerges when information is selected, interpreted and converted into knowledge that can guide action. Artificial intelligence (AI) can therefore be understood not merely as a technical layer, but as one component of a broader knowledge infrastructure through which organisations learn, anticipate and coordinate decisions.

The paper examines the Noverim case as a predictive architecture applied to sustainability ratings and corporate equilibrium assessment. The case is treated as an analytical hypothesis for the IFKAD 2026 debate: if sustainability-related knowledge, organisational signals and predictive analytics are integrated within a coherent rating architecture, firms may be better able to identify emerging disequilibrium before it appears in conventional financial indicators. The reference to 500 Italian knowledge-intensive firms defines only a prospective field of investigation associated with the Noverim framework. It must not be read as empirical evidence, as a validated sample or as the basis of statistical results.

This framing reflects a broader development in management research. Sustainability is no longer only a matter of external legitimacy, disclosure or compliance; it is increasingly recognised as a condition that

influences resilience, stakeholder relationships, innovation capacity and long-term value creation. For knowledge-intensive firms, this point is especially important because value creation depends on intangible assets that are difficult to observe directly but may deteriorate before financial results reveal the decline.

The guiding research question is therefore the following: how can AI-enabled sustainability ratings support early diagnosis, organisational resilience and knowledge-based value creation in knowledge-intensive firms? The paper offers three contributions. Theoretically, it links the knowledge-based view, sustainability management and predictive analytics. Methodologically, it shows how supervised and unsupervised machine-learning approaches may be combined to classify risk and identify trajectories of imbalance. Practically, it advances the notion of a resilience map, in which firms are interpreted through the dynamic interaction of economic, financial, organisational and sustainability dimensions.

2 Theoretical Background

The study is grounded in the view that firms are systems of knowledge, capabilities and relationships. The resource-based view argues that sustained competitive advantage depends on resources that are valuable, rare, imperfectly imitable and organisationally embedded (Barney, 1991; Wernerfelt, 1984). The knowledge-based view extends this argument by treating knowledge as the central productive resource of the firm (Grant, 1996). In knowledge-intensive firms, value creation depends on expertise, problem-solving routines, technological experimentation and the continuous recombination of information. The same features also create vulnerabilities when key knowledge holders leave, learning routines weaken or organisational memory becomes fragmented.

Knowledge management is therefore part of business equilibrium rather than a merely supportive function. Nonaka and Takeuchi (1995) describe knowledge creation as a dynamic process through which tacit and explicit knowledge are continuously converted. Cohen and Levinthal (1990) show that absorptive capacity depends on the ability to recognise, assimilate and exploit external knowledge. These insights are central to resilience: a firm may possess abundant data and still fail to respond if learning routines, governance structures and interpretive capacities are weak.

Sustainability management adds a relational and institutional dimension. Stakeholder theory argues that firms create and preserve value through relationships with multiple constituencies rather than through shareholder interests alone (Freeman, 1984). Corporate sustainability research has shown that environmental, social and governance practices can influence organisational processes, risk exposure and financial performance (Eccles et al., 2014; Maas et al., 2016). This is particularly relevant in knowledge-intensive sectors, where reputation, human capital, trust and legitimacy affect the ability to attract clients, partners and employees.

Predictive analytics shifts the analytical focus from retrospective reporting to forward-looking diagnosis. Prediction is a distinct and valuable objective in information systems research (Shmueli and Koppius, 2011). In management contexts, predictive models can reveal combinations of variables that are difficult to detect through linear reasoning or conventional performance dashboards. Their value, however, depends on interpretability, data quality, validation procedures and managerial judgement. Machine learning should support inquiry and decision-making; it should not replace managerial responsibility.

2.1 Intelligent and sustainable organisations

An intelligent organisation is capable of sensing change, integrating knowledge and adapting behaviour. Intelligence is not reducible to automation; it includes human intelligence, collective intelligence, organisational memory and the selective use of AI. Sustainable organisations pursue economic viability while preserving the social and environmental conditions required for long-term value creation. The intersection of these ideas suggests that sustainability requires intelligence, and intelligence requires sustainability. A technologically advanced firm may remain fragile if it ignores stakeholder trust, employee well-being, governance quality or the ecological and social consequences of its operations.

Hybrid and smart workplaces intensify this connection. Distributed work, digital platforms and collaborative technologies have made knowledge more mobile, but also harder to coordinate and institutionalise. In knowledge-intensive firms, the quality of internal communication, the continuity of learning and the retention of talent become decisive components of resilience. Indicators such as turnover among key people, delays in innovation projects, weak knowledge-sharing routines or reduced investment in training may reveal emerging disequilibrium earlier than conventional financial ratios.

2.2 Corporate equilibrium and early-warning systems

Corporate equilibrium may be understood as the dynamic alignment of profitability, liquidity, capital structure, operational capacity, governance and strategic coherence. Traditional early-warning models made important contributions to bankruptcy prediction through financial ratios and statistical techniques (Altman, 1968; Ohlson, 1980). These models remain useful, but they may be less sensitive to the early phases of decline in firms whose assets are largely intangible and whose vulnerabilities are organisational before they become financial.

A multidimensional equilibrium approach broadens the diagnostic perimeter. It asks whether the firm is economically viable, financially balanced, organisationally coherent and sustainable in its relationships with stakeholders. This view is consistent with integrated risk management (Gatzert and Martin, 2015; Trucco, 2015) and balanced performance measurement (Kaplan and Norton, 1996). The distinctive element of the Noverim case lies in connecting this broader diagnostic logic with predictive intelligence and sustainability ratings.

3 Conceptual Design and Case-Based Analytical Setting

The paper adopts an interpretive, case-based conceptual design centred on the Noverim sustainability-equilibrium model. The aim is not to present a validated predictive formula or original empirical results, but to develop a theoretically grounded study hypothesis on how an intelligent rating architecture could support early diagnosis and resilience-building in knowledge-intensive firms. The case is suitable for this purpose because it connects three dimensions that are often examined separately: corporate crisis prediction, sustainability assessment and knowledge-based organisational resilience.

The analytical setting is defined with reference to a hypothesised group of 500 Italian firms operating in high cognitive-intensity sectors, including advanced professional services, information and communication technology, consulting, applied research, industrial design and high-technology manufacturing. This figure is part of the proposed study design and identifies the potential empirical field on which the Noverim framework could be validated in future research. It does not constitute an observed dataset, an analysed sample or empirical evidence within the present article. Such firms typically depend on qualified human capital, intangible processes and rapid innovation cycles. Their competitive advantage is embedded in people, routines and relationships, but they may also face demand volatility, project-based cash flows, talent dependence and difficulties in translating growth into stable organisational structures.

Methodological clarification: throughout the paper, the expression '500 Italian knowledge-intensive firms' denotes a proposed empirical population for a future validation study. No firm-level observations from these 500 firms are analysed here, and no empirical findings are inferred from them. The present contribution is therefore conceptual and hypothesis-generating, not empirical.

The analysis develops the conceptual structure reported in the extended abstract into a full academic paper. For this reason, all statements concerning predictive capacity, anticipation horizons and crisis patterns are formulated as study hypotheses or analytical propositions. The Noverim case is used as a conceptual and practice-oriented reference, rather than as empirical evidence of model performance. Any future use of the 500-firm reference would require the explicit construction of a dataset, the definition of sampling criteria, the specification of an observation period and the application of appropriate validation metrics. Detailed statistical metrics, calibration procedures, external validation and robustness tests should be developed in a subsequent empirical study if firm-level data become available.

3.1 The Noverim sustainability-equilibrium framework

The Noverim model rests on the assumption that firm viability depends on the interaction of economic, balance-sheet, financial, organisational and sustainability-related conditions. Economic equilibrium concerns margins and productivity. Balance-sheet equilibrium concerns the composition and stability of assets and liabilities. Financial equilibrium relates to liquidity and cash-flow adequacy. Organisational equilibrium concerns governance, internal control and decision quality. Sustainability equilibrium refers to the ways in which environmental, social and governance factors affect risk exposure, adaptive capacity and long-term value creation.

The framework avoids a narrow interpretation of sustainability. In many firms, sustainability information is produced mainly for external reporting, legitimacy or reputational purposes. In the Noverim approach, sustainability is treated as an endogenous determinant of equilibrium. Poor workforce stability, weak governance, inconsistent stakeholder communication or declining investment in capabilities are not peripheral issues. They can alter the firm's risk profile and reduce its capacity to withstand shocks.

Table 1 summarises the main equilibrium dimensions considered in the model. The table should not be read as a mechanical checklist, but as a conceptual map showing how different forms of organisational knowledge can be connected within a predictive architecture. The central diagnostic question is whether signals across dimensions move in a coherent or contradictory direction.

Table 1: Equilibrium dimensions and indicative early-warning signals

Dimension	Knowledge focus	Indicative warning signals
Economic	Margins, productivity and operating continuity	Declining profitability and weaker revenue quality
Balance-sheet	Asset-liability structure and capital solidity	Excessive leverage and fragile capital buffers
Financial	Liquidity, cash-flow timing and debt-service capacity	Cash-flow pressure and dependence on short-term financing
Organisational	Governance, control and decision quality	Key personnel turnover, weak controls and slow response
Sustainability	ESG practices, stakeholder trust and strategic coherence	Deteriorating indicators and strategy-conduct gap

3.2 Predictive architecture

The predictive architecture combines supervised and unsupervised machine-learning techniques. Supervised learning estimates the probability of disequilibrium and improves classification under non-linear conditions. Relevant methods include regression-based approaches, decision-tree models and neural networks. Decision-tree models are useful because they capture interactions among variables and offer a relatively interpretable structure. Neural networks may improve pattern detection in complex data, although their use requires specific attention to transparency, validation and managerial explainability.

Unsupervised learning performs a complementary function. Behavioural clustering detects latent structures among firms and segments them according to similarity in risk profiles and imbalance trajectories. This is particularly useful in knowledge-intensive sectors because crisis pathways are not uniform. Some firms experience liquidity stress after rapid growth; others deteriorate through strategic drift, governance weaknesses or erosion of intangible capabilities.

A key feature of the system is the integration of heterogeneous data at scale. Traditional quantitative indicators are combined with sustainability ratings, organisational variables, qualitative weak signals and information derived from unstructured sources. This extension of the analytical perimeter is valuable, but it also increases the need for data governance, reliability assessment and interpretive discipline.

4 Study Hypotheses and Analytical Propositions

Because the article does not contain original empirical analysis, this section does not report findings in the strict sense. It formulates three study hypotheses derived from the Noverim case and from the theoretical framework. First, the integration of sustainability and organisational indicators with conventional financial measures may improve the anticipation of disequilibrium. Second, knowledge-intensive firms may display recurring pre-crisis patterns that precede visible accounting deterioration. Third, a multidimensional rating architecture may support a meaningful distinction between reversible and structural crises.

4.1 Anticipating disequilibrium

The first study hypothesis is that the Noverim model could anticipate potential imbalance scenarios over an average horizon of 12 to 18 months. This proposition should be treated as an analytical assumption requiring empirical validation, not as a result established by the present article. If confirmed, it would shift the function of rating systems from ex post evaluation to early diagnosis. Conventional financial indicators would remain necessary, but they would become more informative when interpreted together with organisational and sustainability-related signals. A liquidity decline, for example, may be more critical if it coincides with reduced training, increasing turnover and weaker governance quality.

Anticipation should not be confused with deterministic forecasting. A predictive signal indicates an increased probability of disequilibrium, not the certainty of a future event. Its practical value depends on intervention. If managers use the signal to revise strategy, strengthen governance or rebuild knowledge capabilities, the predicted crisis may not occur. In this sense, the absence of a subsequent crisis may indicate the effectiveness of early warning rather than the failure of prediction.

4.2 Recurring pre-crisis patterns

The second study hypothesis concerns the existence of recurring pre-crisis patterns. A first expected pattern is the progressive contraction of intangible investment, including training, research, digital systems, client knowledge and organisational routines. Such reductions may temporarily protect margins, but they can weaken the firm's capacity to innovate, absorb external knowledge and adapt to changing conditions.

The second pattern is increasing turnover among key personnel. Employee mobility is normal in knowledge sectors, but the loss of critical knowledge holders can damage client relationships, project delivery and internal learning. Turnover becomes a warning signal when it affects people who hold tacit knowledge, coordinate strategic projects or maintain important external relationships.

The third pattern is the deterioration of internal sustainability indicators, including governance quality, employee well-being, environmental discipline, transparency and stakeholder engagement. The fourth pattern is growing misalignment between declared strategy and operational behaviour. Such misalignment suggests that the firm may lack the capabilities, governance structures or knowledge routines required to implement its strategic commitments.

These patterns clarify why sustainability ratings can contribute to risk diagnosis. A single weakness does not necessarily imply disequilibrium. However, the simultaneous presence of weaknesses across intangible investment, key personnel stability, governance and operational coherence suggests an erosion of the firm's knowledge base and adaptive capacity.

4.3 Reversible and structural crises

The third study hypothesis concerns the distinction between reversible and structural crises. Reversible crises may be associated with temporary shocks, managerial inefficiencies or limited imbalances that can be corrected through focused intervention. A firm may experience short-term liquidity pressure or delayed projects while remaining viable if its knowledge base, governance quality and stakeholder relationships remain sound.

Structural crises are characterised by simultaneous deterioration across multiple equilibrium dimensions. They may involve declining margins, fragile liquidity, weak governance, reduced intangible investment, loss of key personnel and poor strategic execution. For knowledge-intensive firms, this condition is particularly critical because the assets that disappear are difficult to rebuild quickly. The model supports managerial judgement by showing whether warning signals are concentrated in a limited area or dispersed across the organisational system.

5 Discussion

The analytical propositions developed in this article contribute to the debate on intelligent and sustainable organisations by framing predictive analytics as a potential knowledge-management device. The model does not merely calculate risk. It organises heterogeneous information in a form that may enable managers to observe relationships among sustainability, financial equilibrium and organisational capability. In this sense, the rating process can be interpreted as a possible mechanism of organisational learning.

The case also supports a relational view of sustainability. ESG indicators are not only signals for investors or regulators. They can reveal the quality of governance, the stability of human capital, the credibility of strategy and the capacity to maintain trust. This is important in knowledge-intensive firms, where value is co-produced with employees, clients, partners and communities.

The relationship between human and artificial intelligence is central. AI can improve pattern recognition, but the interpretation of organisational disequilibrium remains a human responsibility. Managers must assess whether model outputs are plausible, ethically acceptable and actionable. From a human-centred perspective, predictive intelligence should support responsible action rather than automate judgement.

The Noverim case also has ecosystem relevance. Knowledge-intensive firms operate within social, business and urban ecosystems where resilience is interdependent. Intelligent rating systems may support not only individual firm management, but also inclusive growth by identifying vulnerabilities, directing advisory support and encouraging preventive intervention before crisis becomes irreversible.

A further issue concerns explainability and organisational acceptance. Managers are more likely to use predictive ratings when they understand the main drivers of a classification and can connect them to lived organisational experience. Feature importance, scenario analysis and structured review meetings can help transform algorithmic output into shared managerial knowledge. Without this interpretive layer, even accurate models may remain external artefacts, consulted for compliance but not embedded in decision-making.

6 Implications for Managers and Policymakers

For managers, the main implication is that sustainability ratings should be used as diagnostic instruments rather than as communication labels. A high score has limited value if it does not inform decisions. A lower score can be useful if it reveals where capabilities need to be strengthened. Managers should monitor the interaction between financial indicators and knowledge-related signals, including training, innovation investment, turnover, client dependence, governance routines and strategic execution.

A second implication concerns measurement culture. Firms with integrated control systems and stronger governance can be hypothesised to be better equipped to prevent disequilibrium. Measurement culture should not be understood as bureaucratic dependence on indicators. It is the disciplined capacity to observe relevant phenomena, discuss signals and act on them. Predictive systems are likely to be useful when they are embedded in such a culture.

For policymakers and ecosystem actors, the case suggests that intelligent sustainability ratings can support preventive approaches to corporate distress. Public agencies, professional associations, financial institutions and innovation intermediaries could use multidimensional early-warning systems to identify firms needing support before insolvency becomes likely. Rating outputs should not become punitive labels, but prompts for advisory, financial and organisational intervention.

For professional practice, the case implies a broader competence profile. Accountants, consultants and managers need to combine financial analysis, sustainability literacy, data interpretation and organisational diagnosis. The future of assessment is unlikely to be purely numerical. It will require the ability to read patterns across multiple domains and to decide whether a signal calls for immediate action, deeper investigation or contextual adjustment.

7 Ethical and Governance Considerations

The use of AI in sustainability ratings raises significant ethical and governance issues. Predictive systems may influence access to credit, investor perceptions, managerial decisions and stakeholder trust. Errors, bias or poor interpretation can therefore produce unfair or disproportionate consequences. Governance should include clear data policies, validation procedures, human oversight and mechanisms for contestability. Firms assessed by such systems should be able to understand the basis of evaluation and correct inaccurate data.

Bias may arise from historical data, incomplete unstructured sources or indicators that reward reporting capacity more than substantive performance. A responsible model should distinguish between absence of evidence and evidence of weakness. It should also avoid over-penalising firms that are transparent about problems while rewarding firms that are merely skilled at disclosure. Human-centred governance is therefore essential to ensure that predictive intelligence supports accountability rather than opacity.

8 Limitations and Future Research

The paper has important limitations. It develops a conceptual argument from the extended abstract and does not report original empirical analysis, statistical metrics, calibration procedures or robustness tests. Consequently, all statements concerning predictive capacity, anticipation horizons, crisis patterns and the 500-firm reference are hypotheses to be tested, not validated results. A future empirical study should define the dataset, sampling criteria, observation period, model specification and validation strategy before any claim of empirical validity is advanced.

Future research could develop longitudinal analyses comparing firms that recover after early-warning signals with those that enter structural decline. Further work should also investigate how managers interpret algorithmic outputs, how explanations affect trust and how predictive systems can support dialogue rather than passive acceptance. Comparative studies across rating architectures would be useful, especially given the divergence in ESG measurement.

9 Conclusions

This paper examined the Noverim case as an analytical hypothesis on how AI-enabled sustainability ratings may support intelligent and sustainable organisations. The argument developed here suggests that corporate equilibrium in knowledge-intensive firms should not be assessed only through conventional accounting indicators. It depends on the interaction of economic performance, financial stability, organisational quality, governance, stakeholder relationships and sustainability-related capabilities.

The central argument is that sustainability is an endogenous component of equilibrium. In knowledge-intensive firms, sustainability indicators often capture the condition of the intangible assets on which value

creation depends: human capital, trust, routines, innovation capacity and strategic coherence. When these assets deteriorate, the firm may become fragile even if short-term financial indicators remain acceptable.

The study hypothesis is that the model may anticipate disequilibrium over a 12-to-18-month horizon and identify patterns such as reduced intangible investment, key personnel turnover, worsening internal sustainability indicators and misalignment between strategy and operations. This claim requires empirical validation before it can be presented as evidence. The same applies to the 500-firm reference, which defines a future field of study, not an empirical basis already analysed in this paper. More broadly, the case suggests that intelligent rating systems can become knowledge infrastructures for resilience-building if they are transparent, ethically governed and embedded in human-centred decision processes.

10 Acknowledgements

The authors acknowledge the contribution of Noverim in developing the sustainability-equilibrium framework discussed in this paper. Any remaining errors or interpretive limitations are the responsibility of the authors.

11 Ethics Declaration

The research design discussed in this paper is conceptual and case-based and does not involve medical experimentation or direct intervention on human subjects. If a future empirical study uses confidential firm-level information, personal data or interviews, the authors should verify whether institutional ethical clearance, informed consent or data-processing approval is required and insert the relevant approval reference before submission.

12 AI Declaration

Generative AI assistance was used during the preparation of this draft to support language refinement, structural expansion from the extended abstract and formatting according to the conference template. The authors should review, revise and validate the full text, conceptual claims, references and final wording before submission. The authors remain responsible for the intellectual content, accuracy and integrity of the submitted paper.

References

Altman, E.I. (1968) Financial ratios, discriminant analysis and the prediction of corporate bankruptcy, The Journal of Finance, Vol. 23, No. 4, September, pp. 589-609.

Barney, J. (1991) Firm resources and sustained competitive advantage, Journal of Management, Vol. 17, No. 1, March, pp. 99-120.

Becchetti, L., Ciciretti, R. and Giovannelli, A. (2012) Corporate social responsibility and earnings forecasting uncertainty, Journal of Business Ethics, Vol. 109, No. 1, July, pp. 53-69.

Berg, F., Koelbel, J.F. and Rigobon, R. (2022) Aggregate confusion: The divergence of ESG ratings, Review of Finance, Vol. 26, No. 6, November, pp. 1315-1344.

Bhimani, A. and Langfield-Smith, K. (2007) Structure, formality and the importance of financial and non-financial information in strategy development and implementation, Management Accounting Research, Vol. 18, No. 1, March, pp. 3-31.

Bini, L. and Bellucci, M. (2020) Integrated reporting and decision usefulness, Meditari Accountancy Research, Vol. 28, No. 4, pp. 767-794.

Breiman, L. (2001) Random forests, Machine Learning, Vol. 45, No. 1, October, pp. 5-32.

Carroll, A.B. (1999) Corporate social responsibility: Evolution of a definitional construct, Business & Society, Vol. 38, No. 3, September, pp. 268-295.

Chen, T. and Guestrin, C. (2016) XGBoost: A scalable tree boosting system, Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, pp. 785-794.

Chesbrough, H.W. (2003) Open Innovation: The New Imperative for Creating and Profiting from Technology, Harvard Business School Press, Boston.

Cohen, W.M. and Levinthal, D.A. (1990) Absorptive capacity: A new perspective on learning and innovation, Administrative Science Quarterly, Vol. 35, No. 1, March, pp. 128-152.

Davenport, T.H. and Prusak, L. (1998) Working Knowledge: How Organizations Manage What They Know, Harvard Business School Press, Boston.

Eccles, R.G., Ioannou, I. and Serafeim, G. (2014) The impact of corporate sustainability on organizational processes and performance, Management Science, Vol. 60, No. 11, November, pp. 2835-2857.

Eisenhardt, K.M. and Martin, J.A. (2000) Dynamic capabilities: What are they?, Strategic Management Journal, Vol. 21, No. 10-11, October-November, pp. 1105-1121.

Elkington, J. (1997) Cannibals with Forks: The Triple Bottom Line of 21st Century Business, Capstone, Oxford.
Freeman, R.E. (1984) Strategic Management: A Stakeholder Approach, Pitman, Boston.
Friede, G., Busch, T. and Bassen, A. (2015) ESG and financial performance: Aggregated evidence from more than 2000 empirical studies, Journal of Sustainable Finance & Investment, Vol. 5, No. 4, pp. 210-233.
García-Sánchez, I.-M., Hussain, N., Khan, S.A. and Martínez-Ferrero, J. (2020) Assurance of sustainability reports and firm performance, Business Strategy and the Environment, Vol. 29, No. 8, December, pp. 3330-3347.
Gatzert, N. and Martin, M. (2015) Determinants and value of enterprise risk management: Empirical evidence from the literature, Risk Management and Insurance Review, Vol. 18, No. 1, Spring, pp. 29-53.
Grant, R.M. (1996) Toward a knowledge-based theory of the firm, Strategic Management Journal, Vol. 17, Winter Special Issue, pp. 109-122.
Hart, S.L. (1995) A natural-resource-based view of the firm, Academy of Management Review, Vol. 20, No. 4, October, pp. 986-1014.
Lundberg, S.M. and Lee, S.-I. (2017) A unified approach to interpreting model predictions, Advances in Neural Information Processing Systems, Vol. 30, pp. 4765-4774.
Hastie, T., Tibshirani, R. and Friedman, J. (2009) The Elements of Statistical Learning: Data Mining, Inference, and Prediction, 2nd ed., Springer, New York.
Kaplan, R.S. and Norton, D.P. (1996) The Balanced Scorecard: Translating Strategy into Action, Harvard Business School Press, Boston.
Maas, K., Schaltegger, S. and Crutzen, N. (2016) Advancing the integration of corporate sustainability measurement, management and reporting, Journal of Cleaner Production, Vol. 133, October, pp. 859-873.
March, J.G. (1991) Exploration and exploitation in organizational learning, Organization Science, Vol. 2, No. 1, February, pp. 71-87.
Nonaka, I. and Takeuchi, H. (1995) The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation, Oxford University Press, New York.
Ohlson, J.A. (1980) Financial ratios and the probabilistic prediction of bankruptcy, Journal of Accounting Research, Vol. 18, No. 1, Spring, pp. 109-131.
Orlitzky, M., Schmidt, F.L. and Rynes, S.L. (2003) Corporate social and financial performance: A meta-analysis, Organization Studies, Vol. 24, No. 3, March, pp. 403-441.
Porter, M.E. and Kramer, M.R. (2011) Creating shared value, Harvard Business Review, Vol. 89, No. 1-2, January-February, pp. 62-77.
Ribeiro, M.T., Singh, S. and Guestrin, C. (2016) Why should I trust you? Explaining the predictions of any classifier, Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, pp. 1135-1144.
Shmueli, G. and Koppius, O.R. (2011) Predictive analytics in information systems research, MIS Quarterly, Vol. 35, No. 3, September, pp. 553-572.
Teece, D.J. (2007) Explicating dynamic capabilities: The nature and microfoundations of sustainable enterprise performance, Strategic Management Journal, Vol. 28, No. 13, December, pp. 1319-1350.
Trucco, P. (2015) Integrated Risk Management: Theoretical Framework and Practical Applications, Springer, Heidelberg.
Wernerfelt, B. (1984) A resource-based view of the firm, Strategic Management Journal, Vol. 5, No. 2, April-June, pp. 171-180.
Waddock, S.A. and Graves, S.B. (1997) The corporate social performance-financial performance link, Strategic Management Journal, Vol. 18, No. 4, April, pp. 303-319.
Zopounidis, C. and Doumpos, M. (2002) Multicriteria decision aid in financial decision making, European Journal of Operational Research, Vol. 136, No. 2, January, pp. 361-379.

Intelligent Knowledge, Corporate Crisis and Criminal Liability: Incentive Structures for Sustainable Governance in the Italian Framework

Alfonso Laudonia
Universitas Mercatorum, Rome, Italy

Abstract

The paper investigates how intelligent knowledge systems reshape incentives and responsibilities in corporate crisis governance under the Italian legal framework. It focuses on the interaction between the Code of Business Crisis and Insolvency (CCII) and Legislative Decree 231/2001 on corporate criminal liability, analysing how legal duties of early detection and adequate organisational arrangements intersect with the emerging use of AI-enabled early-warning and monitoring tools. The central question is under which conditions these systems function as levers for sustainable and responsible governance, and when they instead generate new risks of moral hazard, scapegoating and purely formalistic compliance.

Theoretically, the paper bridges three strands of literature: knowledge management and organisational learning, which conceptualise knowledge as a strategic asset increasingly mediated by digital technologies (Nonaka and Takeuchi, 1995; Argote, 2013; Schiuma, 2011); law and economics of corporate crisis and insolvency, which examines how sanctions and incentives shape directors' behaviour in the vicinity of insolvency (Jensen and Meckling, 1976; Coffee, 2006; Armour and Ringe, 2020); and socio-legal and criminological studies on corporate crime and compliance, which highlight the role of organisational cultures and control architectures in enabling or constraining wrongdoing (Clinard and Yeager, 1980; Braithwaite, 1989; Parker, 2002). Building on these perspectives, the paper proposes to conceptualise corporate crisis governance as a knowledge-based regulatory game in which legal rules, intelligent knowledge systems and organisational cultures jointly determine the information and payoff structures faced by corporate actors.

Methodologically, the study adopts a qualitative, analytical-conceptual design that combines doctrinal analysis of the CCII and Legislative Decree 231/2001 with theory-informed discussion of AI-enabled early-warning and HR-analytics-based frameworks, including prior empirical work by the author on predictive HR analytics for financial risk prevention (Laudonia et al., 2025).

The analysis shows that intelligently designed and well-governed knowledge systems can enhance the preventive aims of the CCII and the compliance-oriented incentives of the 231 regime by reducing information asymmetries, supporting timely escalation and enabling more informed restructuring decisions. However, it also warns against the risks of black-box solutions that may foster over-reliance on algorithms, diffusion of responsibility and symbolic compliance. The paper concludes by outlining design principles for intelligent knowledge systems in crisis governance—transparency, explainability, preservation of professional judgement, robust data-governance and learning-oriented feedback loops—and by deriving implications for regulators, boards and compliance officers concerned with aligning legal accountability, effective risk reduction and long-term organisational sustainability.

Keywords: intelligent knowledge systems; corporate crisis governance; early-warning; criminal liability; sustainable governance.

Paper type: Academic Research Paper

1 Introduction

1.1 Background and problem statement

The convergence between corporate crisis management, corporate criminal liability and knowledge management is reshaping how organisations govern financial distress in complex and uncertain environments. In many jurisdictions, and particularly in Italy, recent legal reforms have strengthened directors' duties to detect and address business crisis at an early stage, while expanding corporate and quasi-criminal liability for organisational failures under regimes such as the Codice della crisi d'impresa e dell'insolvenza (CCII) and Legislative Decree 231/2001. At the same time, advances in artificial intelligence (AI), data analytics and digital infrastructures have enabled new forms of "intelligent knowledge" systems that integrate human expertise, organisational routines and algorithmic tools in monitoring, decision-making and control processes (Nonaka and Takeuchi, 1995; Arrieta et al., 2020; Rame, 2024). Yet it remains unclear under which conditions these systems actually enhance sustainable and responsible governance, and when instead they generate new risks of moral

hazard, formalistic compliance or diffusion of responsibility among corporate actors (Laudonia et al., 2025; Gilad, 2010).

1.2 Research gap and research questions

The knowledge management and organisational learning literature has extensively analysed how knowledge assets, learning routines and digital technologies contribute to organisational performance and resilience (Nonaka and Takeuchi, 1995; Argote, 2013; Schiuma, 2011), but has paid limited attention to the specific context of corporate crisis governance and to the interaction with legal liability regimes. In parallel, the law and economics literature on corporate crisis and insolvency has focused on sanctions, incentives and early-warning mechanisms, often modelling directors' behaviour under conditions of information asymmetry and moral hazard (Coffee, 2006; Armour and Ringe, 2020), yet it has rarely incorporated the role of knowledge flows, AI-based analytics and organisational learning. Socio-legal and criminological studies on corporate crime and compliance have examined organisational cultures, control systems and "tone at the top" (Clinard and Yeager, 1980; Braithwaite, 1989), but they have not systematically integrated the emerging class of AI-enabled early-warning and monitoring tools into their analyses (Arrieta et al., 2020; Parker, 2002).

As a result, there is still no integrated framework that conceptualises corporate crisis governance as a knowledge-based regulatory game in which legal rules, intelligent knowledge systems and organisational cultures jointly determine the incentives, perceived risks and actual choices of corporate actors. This paper addresses this gap by asking two research questions: RQ1: How do intelligent knowledge systems, including AI-based early-warning and monitoring tools, interact with the Italian legal framework on business crisis and corporate criminal liability, particularly the CCII and Legislative Decree 231/2001? RQ2: Under what conditions do these systems function as levers for sustainable and responsible governance, rather than as catalysts of new risks and dysfunctional behaviours?

1.3 Purpose and contributions of the paper

The purpose of the paper is to develop a conceptual model of corporate crisis governance that explicitly integrates legal-institutional arrangements, intelligent knowledge systems and incentive structures. Building on knowledge management (Nonaka and Takeuchi, 1995; Schiuma, 2011), law and economics of corporate governance and insolvency (Coffee, 2006; Armour and Ringe, 2020), and corporate crime and compliance studies (Braithwaite, 1989; Parker, 2002), the paper offers three main contributions. First, it bridges these strands of literature by articulating how knowledge flows and AI-mediated analytics interact with sanctions, rewards and organisational cultures in the context of business crisis. Second, it refines the analysis of sanctions and incentives in corporate and insolvency law by incorporating dimensions of knowledge quality, organisational learning and technological mediation, drawing on empirical evidence about AI-driven early-warning and HR analytics in crisis prevention (Laudonia et al., 2025; Altman, 1968). Third, it proposes design principles for intelligent knowledge systems in crisis governance that aim to align legal compliance, effective risk reduction and long-term organisational sustainability.

1.4 Structure of the paper

The remainder of the paper is structured as follows. Section 2 reviews the main strands of literature and synthesises the theoretical gap. Section 3 presents the Italian legal and institutional context. Section 4 outlines the methodology. Section 5 discusses intelligent knowledge systems in corporate crisis governance, with particular attention to AI-enabled early-warning tools. Section 6 develops the core analysis on incentives, liability and governance outcomes. Section 7 derives design principles and policy implications, and Section 8 concludes.

2 Theoretical Background and Literature Review

2.1 Knowledge management, intelligent knowledge and organisational learning

Knowledge management (KM) literature has long conceptualised knowledge as a strategic resource embedded in organisational routines, processes and social relationships (Nonaka and Takeuchi, 1995; Argote, 2013). Dynamic perspectives on organisational learning emphasise how firms develop and renew knowledge-based capabilities through exploration, exploitation and feedback loops between individual and collective levels (March, 1991; Crossan et al., 1999). More recent work has highlighted the role of digital technologies and analytics in enabling new forms of "intelligent" knowledge systems that combine human cognition with

algorithmic processing and data-driven insights (Schiuma, 2011; Rame, 2024). In this view, intelligent knowledge emerges from the interaction between human judgement, data architectures and AI-enabled tools that support interpretation, prediction and decision-making under uncertainty (Arrieta et al., 2020; Witten et al., 2016). However, this stream has rarely engaged with legal constraints and liability regimes that shape how knowledge is produced, shared and used in high-stakes contexts such as corporate crisis governance.

2.2 Law and economics of corporate crisis and insolvency

Law and economics scholarship explains corporate crisis and insolvency primarily through incentives, information asymmetries and moral hazard. In this perspective, legal rules on early warning, directors' duties and restructuring are designed to encourage timely intervention and discourage opportunistic conduct (Jensen and Meckling, 1976; Coffee, 2006; Armour and Ringe, 2020; Eidenmüller, 2017). However, this literature has paid less attention to the organisational and technological infrastructures through which relevant knowledge is generated, processed and acted upon, especially in the context of AI-based monitoring and prediction.

2.3 Corporate crime, organisational wrongdoing and compliance architectures

Socio-legal and criminological research has shown that organisational misconduct is shaped not only by formal rules, but also by cultures, control systems and internal compliance architectures (Clinard and Yeager, 1980; Braithwaite, 1989; Parker, 2002; Gilad, 2010). More recent studies on AI governance and corporate compliance suggest that algorithmic monitoring tools may affect accountability, auditability and oversight, but they have not yet systematically integrated AI-enabled early-warning systems into broader models of corporate wrongdoing and compliance.

2.4 Synthesis and identified literature gap

Taken together, these strands of literature show an important gap. Knowledge management research explains how intelligent systems support organisational adaptation, law and economics clarifies how legal rules shape incentives in crisis, and corporate crime scholarship highlights the role of culture and compliance architectures; yet none of these perspectives, taken alone, offers an integrated account of how legal rules, intelligent knowledge systems and organisational cultures jointly shape behaviour in corporate distress. This paper addresses that gap by conceptualising corporate crisis governance as a knowledge-based regulatory game, with specific reference to the interaction between the Italian CCII, Legislative Decree 231/2001 and AI-enabled early-warning systems.

3 Legal and Institutional Context: the Italian Framework

3.1 The Italian Code of Business Crisis and Insolvency (CCII)

The Italian Code of Business Crisis and Insolvency (CCII) marks a shift from a liquidation-centred approach to a preventive model of crisis governance. It requires directors to adopt adequate organisational, administrative and accounting arrangements capable of detecting early signs of distress and supporting timely intervention, thereby making information quality and internal monitoring central to corporate crisis management (Armour and Ringe, 2020).

3.2 Corporate criminal liability and organisational models under Legislative Decree 231/2001

Legislative Decree 231/2001 establishes corporate liability for offences committed in the entity's interest or to its advantage, unless the firm can demonstrate the adoption and effective implementation of adequate organisation, management and control models. From this perspective, 231 models are not only compliance documents, but governance architectures based on procedures, information flows, risk mapping and oversight mechanisms, increasingly supported by digital reporting and monitoring tools (Coffee, 2006; Parker, 2002).

3.3 Recent reforms and debates on crisis, liability and sanctions

Recent Italian debates on crisis governance and sanctions have reinforced attention to directors' responsibilities, the calibration of deterrence, and the coordination between insolvency, corporate and criminal law, while also raising new questions about how AI-based monitoring tools should be assessed in terms of adequacy and accountability (Eidenmüller, 2017; Arrieta et al., 2020; Laudonia et al., 2025).

3.4 Intelligent knowledge and governance in the Italian context

Taken together, the CCII and Legislative Decree 231/2001 make the Italian framework a particularly relevant setting for analysing intelligent knowledge systems in corporate crisis governance. Both regimes depend, implicitly or explicitly, on the organisation's ability to generate, process and use reliable knowledge about risk, distress and misconduct; for this reason, the Italian case is especially useful for examining how AI-enabled early-warning systems may strengthen prevention and compliance, but also create new risks of opacity, formalism and responsibility-shifting.

4 Methodology

4.1 Overall research design

The paper adopts a qualitative, analytical-conceptual design that combines doctrinal legal analysis with theory-informed examination of organisational and technological developments (Eisenhardt, 1989; Yin, 2018). Rather than testing specific hypotheses on a large-N dataset, it aims to build an integrated framework linking legal-institutional arrangements, intelligent knowledge systems and incentive structures in corporate crisis governance.

4.2 Sources and data

The analysis relies on three main categories of sources. First, primary legal materials, including the CCII and Legislative Decree 231/2001, are used to reconstruct duties, sanctions and institutional mechanisms shaping corporate crisis governance (Armour and Ringe, 2020). Second, secondary legal and socio-legal literature on corporate governance, insolvency, corporate crime and compliance informs the main debates on incentives and liability (Coffee, 2006; Parker, 2002; Gilad, 2010). Third, empirical and conceptual studies on AI-enabled early-warning systems and HR-analytics-based crisis prediction provide evidence on how intelligent knowledge systems operate in practice and affect information flows and decision-making under distress, including previous co-authored work on AI-driven HR analytics and financial risk prevention (Laudonia et al., 2025; Altman, 1968).

4.3 Analytical strategy

Analytically, the paper proceeds in three steps. First, it maps how the CCII and Legislative Decree 231/2001 allocate responsibilities, define "adequate" organisational and control arrangements, and link them to corporate and personal liability (Armour and Ringe, 2020; Coffee, 2006). Second, it overlays this legal structure with insights from knowledge management and organisational learning to conceptualise the underlying knowledge infrastructures and identify where AI-based tools are most likely to be embedded in crisis governance processes (Nonaka and Takeuchi, 1995; Argote, 2013; Schiuma, 2011). Third, it uses selected examples from AI-enabled early-warning and HR-analytics systems, together with law-and-economics reasoning, to articulate how different configurations of legal rules, knowledge systems and organisational cultures generate distinct incentive patterns (Laudonia et al., 2025; Gilad, 2010).

4.4 Limitations

This approach has the usual limitations of analytical-conceptual and case-informed research designs (Eisenhardt, 1989; Yin, 2018). The study does not provide large-scale empirical testing or quantitative estimates of the impact of intelligent knowledge systems on crisis outcomes or liability exposure, and its empirical base relies on existing studies and doctrinal interpretation rather than original fieldwork (Laudonia et al., 2025). The focus on the Italian context may also constrain direct generalisation to other legal systems with different corporate and insolvency law architectures (Armour and Ringe, 2020), although the framework is intended to support subsequent empirical and comparative research (Eidenmüller, 2017).

5 Intelligent Knowledge Systems in Corporate Crisis Governance

5.1 From traditional early-warning systems to intelligent knowledge

Traditional early-warning systems (EWS) for corporate crises have relied mainly on financial ratios, balance-sheet indicators and static thresholds to signal distress (Altman, 1968; Armour and Ringe, 2020). These tools capture deteriorating solvency or liquidity, but struggle to anticipate crises driven by human capital dynamics,

organisational misalignment or slow capability erosion (Rame, 2024). In more complex Industry 5.0 environments, early-warning therefore needs to integrate heterogeneous data sources, including HR indicators, operational signals and qualitative assessments of organisational climate. Intelligent knowledge systems extend traditional EWS by combining structured and unstructured data, using machine-learning and natural language processing to detect latent risk patterns, and embedding these outputs into governance processes where human decision-makers can interpret and act on them (Witten et al., 2016; Arrieta et al., 2020).

5.2 AI-based early-warning and HR analytics as a paradigmatic case

AI-driven HR analytics provide a paradigmatic example of intelligent knowledge systems for crisis governance. Laudonia et al. (2025) develop a framework in which HR variables—such as turnover, skill gaps, training investments and sentiment from performance reviews—are combined with financial indicators to predict financial distress. Their results suggest that predictive HR analytics can reduce false alarms and identify a substantial share of crises in advance, improving the sensitivity and precision of early-warning signals (Laudonia et al., 2025). These frameworks rely on integrating HR and financial data with text-mining applied to qualitative sources, using machine-learning models to detect patterns associated with subsequent downturns (Witten et al., 2016), while supporting rather than replacing managerial judgement in strategic and governance decisions.

5.3 Embedding intelligent knowledge systems into CCII and 231 governance architectures

In the Italian framework, intelligent knowledge systems intersect directly with duties and incentives under the CCII and Legislative Decree 231/2001. Under the CCII, advanced predictive analytics can form part of the "adequate organisational, administrative and accounting arrangements" required to detect early signs of crisis and support timely intervention, as long as they remain transparent, reliable and aligned with decision-making processes (Armour and Ringe, 2020; Laudonia et al., 2025). Within the 231 regime, organisation, management and control models are assessed for their adequacy to prevent offences and ensure effective risk monitoring; intelligent systems that integrate HR, operational and financial data can enhance these models by enabling granular risk mapping, continuous anomaly detection and evidence-based oversight (Coffee, 2006; Parker, 2002). At the same time, complex AI-driven tools raise concerns about accountability, explainability and "black box compliance", where formal reliance on sophisticated systems masks weaknesses in governance or culture and complicates the assessment of directors' and entities' liability (Arrieta et al., 2020; Eidenmüller, 2017; Gilad, 2010).

6 Analysis and Discussion: Incentives, Liability and Governance Outcomes

6.1 Intelligent knowledge as a lever for sustainable and responsible governance

When intelligently designed and properly embedded, knowledge systems can strengthen the preventive and sustainability-oriented aims of corporate crisis governance. By integrating financial, HR and operational indicators into robust early-warning architectures, they help boards and managers detect weak signals of distress, evaluate options and initiate timely restructuring or negotiation with creditors (Altman, 1968; Armour and Ringe, 2020). This supports the policy objectives of the CCII and European preventive restructuring frameworks and, from an incentive perspective, modifies directors' perceived payoff structure: credible, data-driven tools increase the expected costs of concealment and inaction relative to early disclosure and proactive engagement (Coffee, 2006; Armour and Ringe, 2020; European Commission, 2025). When oversight bodies have access to granular, up-to-date information, they are also better placed to exercise effective monitoring and counter the normalisation of deviance (Braithwaite, 1989; Vaughan, 1999).

6.2 Risks of moral hazard, scapegoating and formalistic compliance

If poorly designed or weakly governed, AI-enabled early-warning and monitoring tools can generate new forms of moral hazard and opportunities for scapegoating. Managers may treat sophisticated analytics as substitutes rather than complements to their own vigilance, leading to over-reliance on algorithmic outputs and under-investment in qualitative assessment (Arrieta et al., 2020). Boards and oversight bodies may invoke the mere existence of advanced systems as evidence of "adequate" arrangements, even when those systems are poorly calibrated, under-resourced or routinely ignored, thereby reinforcing patterns of window-dressing and paper compliance already documented in the corporate crime literature (Parker, 2002; Laudonia et al., 2025). In

such scenarios, intelligent knowledge systems risk becoming shields rather than levers for accountability, complicating the attribution of fault under regimes such as the CCII and Legislative Decree 231/2001 (Coffee, 2006; Eidenmüller, 2017).

6.3 Corporate crisis governance as a knowledge-based regulatory game

The interplay between legal rules, intelligent knowledge systems and organisational cultures can be conceptualised as a knowledge-based regulatory game. The main players—directors, managers, oversight bodies, regulators and, indirectly, creditors and employees—interact within legal and organisational rules that define duties, sanctions, rewards and information rights (Jensen and Meckling, 1976; Coffee, 2006). The state of the game depends on the quality and distribution of knowledge about the firm's condition, the configuration of intelligent knowledge systems (their scope, transparency and governance), and the expectations that players hold about how others will react to distress signals. Intelligent systems thus reshape both the information and the payoff structure of the game, influencing whether early disclosure of crisis is perceived as a dominated strategy or as a potentially advantageous move under regimes such as the CCII and 231 (Witten et al., 2016; Laudonia et al., 2025; Armour and Ringe, 2020).

6.4 Implications for sanctions, enforcement and organisational learning

Viewing corporate crisis governance as a knowledge-based regulatory game has several implications for sanctions and enforcement. Liability assessments should consider not only the formal existence of intelligent knowledge systems, but also their substantive functioning, including responsibility for their design and maintenance, documentation of outputs and actual responses to early-warning signals (Gilad, 2010; Laudonia et al., 2025). Enforcement strategies could be calibrated to reward transparent, explainable and well-governed systems and to penalise purely symbolic or black-box implementations (Arrieta et al., 2020; European Commission, 2025). Finally, crisis events and enforcement outcomes can themselves become sources of organisational learning, fostering a co-evolution of intelligent knowledge and legal incentives towards more sustainable and responsible corporate governance (Nonaka and Takeuchi, 1995; Schiuma, 2011).

7 Design Principles and Policy Implications

7.1 Design principles for intelligent knowledge systems in crisis governance

A first implication of the analysis is the need to design intelligent knowledge systems according to principles that align legal compliance, effective risk reduction and organisational sustainability. Transparency and explainability are essential: AI-enabled tools used for early-warning and monitoring should provide interpretable outputs and clear decision-rules so that directors, oversight bodies and regulators can understand how signals are generated and on what basis they should act (Arrieta et al., 2020; Witten et al., 2016). A second principle is the preservation and valorisation of professional judgement, ensuring that human decision-makers retain ultimate responsibility and that algorithms are used to augment, not replace, critical assessment (Nonaka and Takeuchi, 1995; Laudonia et al., 2025). Third, intelligent knowledge systems should be tightly coupled with organisational routines and escalation procedures, so that early-warning signals systematically trigger documented discussions, decisions and follow-up actions (Armour and Ringe, 2020; Gilad, 2010). A fourth principle concerns the integration of ethical, data-protection and fairness safeguards, particularly when HR and other sensitive data are involved (European Commission, 2025). Finally, intelligent knowledge systems should incorporate feedback loops that transform crisis experiences and enforcement outcomes into learning opportunities (Schiuma, 2011; Rame, 2024).

7.2 Implications for regulators and policymakers

For regulators and policymakers, the findings suggest that legal frameworks such as the CCII and Legislative Decree 231/2001 should explicitly recognise the role of intelligent knowledge systems while setting expectations regarding their governance. Guidance and soft-law instruments could clarify how AI-enabled early-warning and monitoring tools may contribute to "adequate organisational and control arrangements", and which minimal standards of transparency, documentation and oversight are required for such tools to be considered in liability assessments (Armour and Ringe, 2020; Eidenmüller, 2017). Policymakers might also explore incentive schemes—such as reduced sanctions or procedural advantages—for firms that adopt robust, explainable and independently audited knowledge systems, thereby steering investments towards genuinely effective

infrastructures rather than symbolic compliance (Coffee, 2006; European Commission, 2025). At the same time, regulators should remain attentive to the risks of over-reliance on technology, and supervisory practices may need to develop specialised competences in assessing AI-enabled systems, including the capacity to scrutinise data-governance arrangements and model validation procedures (Arrieta et al., 2020; Gilad, 2010).

7.3 Implications for boards, compliance officers and practitioners

For boards of directors, compliance officers and other practitioners involved in crisis governance, the analysis underscores the strategic nature of intelligent knowledge systems. Boards should treat the design and oversight of these systems as a core governance responsibility, ensuring alignment with the firm's risk appetite, values and legal obligations, and avoiding delegation of critical design choices solely to technical experts or external vendors (Braithwaite, 1989; Parker, 2002). Compliance and risk-management functions should work closely with HR, finance and IT departments to integrate data sources, define meaningful indicators and develop protocols that translate early-warning signals into concrete actions and accountability trails (Laudonia et al., 2025). Practitioners should also be aware of the reputational and legal implications of black-box solutions: investing in explainability, documentation and training for users can not only improve the effectiveness of intelligent knowledge systems, but also demonstrate diligence and good faith in the event of investigations or litigation related to crisis management failures (Arrieta et al., 2020; Gilad, 2010). Ultimately, the practical challenge is to move from viewing intelligent knowledge systems as compliance add-ons to recognising them as central components of a broader governance architecture in which legal rules, organisational culture and technological infrastructures co-evolve.

8 Conclusions

This paper has argued that corporate crisis governance can be fruitfully conceptualised as a knowledge-based regulatory game, in which legal rules, intelligent knowledge systems and organisational cultures jointly shape the incentives, perceived risks and behavioural choices of corporate actors. By integrating insights from knowledge management and organisational learning, law and economics of crisis and insolvency, and socio-legal studies of corporate crime and compliance, the analysis has highlighted how AI-enabled early-warning tools and HR-analytics-based frameworks—such as those proposed by Laudonia et al. (2025)—can transform the information structures underpinning directors' duties and liability regimes under the Italian CCII and Legislative Decree 231/2001 (Nonaka and Takeuchi, 1995; Coffee, 2006; Laudonia et al., 2025).

The main contribution is twofold. Theoretically, the paper offers an integrated framework that foregrounds the role of intelligent knowledge in mediating between legal incentives and organisational behaviour, showing that the effectiveness of preventive and sanctioning mechanisms depends crucially on the design and governance of socio-technical knowledge infrastructures. Practically, it derives design principles and policy implications for boards, compliance officers and regulators, emphasising transparency, explainability, preservation of professional judgement, robust data-governance and the embedding of feedback loops that turn crisis experiences and enforcement outcomes into organisational learning (Arrieta et al., 2020; Schiuma, 2011; European Commission, 2025).

The analysis is not without limitations. It rests on an analytical-conceptual approach and on existing empirical studies, including prior work by the author, rather than on original large-scale data, and it focuses primarily on the Italian context, which may differ significantly from other legal systems (Eidenmüller, 2017; Armour and Ringe, 2020). Future research could extend the framework through comparative studies of different jurisdictions, in-depth case studies of firms implementing intelligent knowledge systems in crisis situations, and quantitative assessments of how such systems affect crisis trajectories, enforcement outcomes and the calibration of sanctions. By pursuing these lines of inquiry, scholars and practitioners can further refine our understanding of how intelligent knowledge can support truly sustainable and responsible corporate governance in times of financial distress.

9 Acknowledgements

The author wishes to thank the IFKAD 2026 reviewers for their constructive comments, which significantly contributed to improving the quality and focus of this paper.

10 Ethics Declaration

This paper does not report empirical research involving human participants, personal data or experiments requiring formal ethics committee approval. All materials discussed are based on publicly available legal sources and secondary literature. No ethical clearance was therefore required for this research.

11 AI Declaration

The author used AI-based tools (Perplexity AI) to support literature searches, language editing and drafting of early versions of some sections of this paper. All ideas, arguments, theoretical framework, citations and final formulations have been critically reviewed, validated and, where necessary, substantially revised by the author, who retains full responsibility for the content and any errors or omissions.

References

Altman, E.I. (1968) "Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy", 'Journal of Finance', Vol. 23, No. 4, pp. 589–609.

Argote, L. (2013) 'Organizational Learning: Creating, Retaining and Transferring Knowledge', 2nd edn, Springer, New York.

Armour, J. and Ringe, W.G. (2020) "Comparative Corporate Insolvency Law", in Armour, J. et al. (eds), 'Oxford Handbook of Corporate Law and Governance', Oxford University Press, Oxford.

Arrieta, A.B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., García, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R. and Herrera, F. (2020) "Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges toward Responsible AI", 'Information Fusion', Vol. 58, pp. 82–115.

Braithwaite, J. (1989) 'Crime, Shame and Reintegration', Cambridge University Press, Cambridge.

Clinard, M.B. and Yeager, P.C. (1980) 'Corporate Crime', Free Press, New York.

Coffee Jr., J.C. (2006) 'Gatekeepers: The Professions and Corporate Governance', Oxford University Press, Oxford.

Crossan, M.M., Lane, H.W. and White, R.E. (1999) "An Organizational Learning Framework: From Intuition to Institution", 'Academy of Management Review', Vol. 24, No. 3, pp. 522–537.

Eidenmüller, H. (2017) "Comparative Corporate Insolvency Law", in Armour, J. and Payne, J. (eds), 'Rationality in Company Law: Essays in Honour of DD Prentice', Hart Publishing, Oxford and Portland.

Eisenhardt, K.M. (1989) "Building Theories from Case Study Research", 'Academy of Management Review', Vol. 14, No. 4, pp. 532–550.

European Commission (2025) 'Directive on Preventive Restructuring Frameworks: Implementation Report', European Commission, Brussels.

Gilad, S. (2010) "Enlisting Commitment to Internal Compliance via Reframing and Delegation", 'LSE CARR Discussion Paper', No. 64, London School of Economics, London.

Jensen, M.C. and Meckling, W.H. (1976) "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure", 'Journal of Financial Economics', Vol. 3, No. 4, pp. 305–360.

Laudonia, A., [Co-authors] (2025) "AI-Driven Financial Risk Prevention: The Role of HR Analytics in Corporate Crisis Management under Industry 5.0", 'Proceedings of the International Conference on Industry of the Future and Smart Manufacturing'.

March, J.G. (1991) "Exploration and Exploitation in Organizational Learning", 'Organization Science', Vol. 2, No. 1, pp. 71–87.

Nonaka, I. and Takeuchi, H. (1995) 'The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation', Oxford University Press, New York.

Parker, C. (2002) 'The Open Corporation: Effective Self-Regulation and Democracy', Cambridge University Press, Cambridge.

Rame, A. (2024) "Intelligent Knowledge and Organisational Systems in the Age of Industry 5.0", 'Journal of Knowledge Management', [forthcoming/in press].

Schiuma, G. (2011) "Managing and Measuring Knowledge Assets Dynamics for Business Value Creation", in Schiuma, G. (ed.), 'Managing Knowledge Assets and Business Value Creation in Organizations: Measures and Dynamics', IGI Global, Hershey, pp. 1–17.

Vaughan, D. (1999) "The Dark Side of Organizations: Mistake, Misconduct, and Disaster", 'Annual Review of Sociology', Vol. 25, pp. 271–305.

Witten, I.H., Frank, E., Hall, M.A. and Pal, C.J. (2016) 'Data Mining: Practical Machine Learning Tools and Techniques', 4th edn, Morgan Kaufmann, Burlington, MA.

Yin, R.K. (2018) 'Case Study Research and Applications: Design and Methods', 6th edn, SAGE, Thousand Oaks, CA.

From Cost to Social Capacity: Governing AI-Released Time with STRR and S-STRR

Barbara Fasoli Braccini, Federico Gaeta
University of Rome UnitelmaSapienza, Rome, Italy

Abstract

The adoption of generative artificial intelligence and automation technologies is reshaping how organisations create and allocate value. While a growing body of evidence shows that AI can generate measurable productivity gains, support workplace learning, and improve selected aspects of job quality, both research and managerial practice still lack an auditable pathway capable of linking AI-enabled time savings to reinvestment decisions and their internal and societal effects.

Starting from this gap, the paper introduces a fundamental distinction between *Created Value* and *Released Time* (RT). Created Value refers to the economic and competitive output generated through organisational activities, whereas Released Time identifies the stock of time, attention, and relational capacity made available by the reduction of routine or low-value tasks. We conceptualise Released Time as a governable resource rather than an automatic source of value.

Building on this foundation, the paper develops an integrated measurement-to-governance architecture. First, Released Time is quantified through Time-Driven Activity-Based Costing (TDABC) time equations, adjusted by a quality–risk–compliance factor to exclude spurious efficiency gains. Second, we introduce the *Social Time Reinvestment Ratio* (STRR), a micro-level governance indicator (0–1) that captures the share of eligible Released Time intentionally reallocated to work-life balance, learning and coordination, and skills-based volunteering. STRR is designed to be auditable through allocation logs, eligibility rules, documentary evidence, and reconciliation with process accounts.

The framework is then extended to the *Societal STRR* (S-STRR), which connects reinvested time portfolios to Social Return on Investment (SROI) logic models and proxy maps, incorporating standard adjustments such as deadweight, attribution, displacement, and drop-off. This enables the identification of a "double dividend": internal human capital strengthening and monetised societal value.

Methodologically, the paper adopts a conceptual-development approach with an explicit early-stage validation protocol. It defines constructs, operational rules, and an evidence framework ensuring reliability (inter-rater agreement on allocation decisions) and integrity (reconciliation between released time and TDABC accounts).

By integrating previously disconnected literatures, namely AI productivity, TDABC, SROI, skills-based volunteering, and AI governance, the paper contributes an auditable measurement-to-governance architecture.

Keywords: Time-Driven Activity-Based Costing; Released Time; Social Time Reinvestment Ratio; Skills-based Volunteering; Social Return on Investment.

Paper type: Academic Research Paper

1 Introduction

Generative artificial intelligence is producing measurable productivity gains in real workplace settings, from controlled experiments to large-scale deployments in knowledge work. Evidence shows that AI can increase efficiency, support learning, and improve aspects of job quality. However, these gains are highly heterogeneous and do not automatically translate into sustainable organisational capability or societal value.

This highlights a critical governance gap in the way organisations manage Released Time. Organisations lack an auditable pathway linking AI-enabled time savings to reinvestment decisions and to their internal and external effects. As a result, Released Time may be dissipated, absorbed into work intensification, or only partially converted into higher-value or socially relevant activities.

Part of this problem reflects the fragmentation of the literature. Research on AI productivity, Time-Driven Activity-Based Costing (TDABC), and Social Return on Investment (SROI) remains largely disconnected, while AI governance studies often struggle to translate principles into operational and measurable practices. We treat Released Time as a governable stock, not value.

By integrating AI productivity, TDABC, SROI, and governance literatures, the paper develops an auditable framework that connects measurement with organisational decision-making.

This perspective is also consistent with the intellectual capital tradition, which conceptualises organisational value as the result of interactions between human, structural, and relational capital (Edvinsson and Malone, 1997). In this view, AI-enabled Released Time can be interpreted as a latent form of multi-capital resource, comprising time, attention, and relational capacity, that may be converted into human capital (through learning and skill development), structural capital (through codification and coordination), and relational capital (through

skills-based volunteering and stakeholder engagement) (Nonaka and Takeuchi, 1995). This positioning reinforces the need to treat Released Time not as a residual outcome of efficiency, but as a governable driver of knowledge-based value creation.

This study adopts a conceptual-development and design-oriented research approach, aimed at constructing an auditable governance artefact rather than testing causal relationships.

2 Why Existing Approaches are Insufficient for Governing Released Time

Current AI productivity research measures output and time savings and increasingly documents heterogeneity and task reallocation. For example, developer studies show both productivity increases and shifts away from non-core activities (Hoffmann et al., 2025), while large-scale field experiments highlight time saved in communication and after-hours work without necessarily changing task composition (Dillon et al., 2025). These findings sharpen the governance gap: time is released, but organisational reinvestment is not automatically observed or governed.

TDABC's strength is to quantify time at activity level through time equations and capacity cost rates, enabling transparent measurement and reconciliation (Etges, Polanczyk and Urman, 2020). Recent reviews stress the expanding role of TDABC and the need for standardised reporting to improve comparability and reduce variability (Etges et al., 2024; Shakya et al., 2025). However, TDABC does not specify where released time goes, nor does it specify governance rules for reinvestment portfolios, making measurement alone insufficient.

SROI frameworks are increasingly refined and reviewed, with advances aimed at improving transparency and comparability (Ashton et al., 2024; Gutiérrez-Nieto et al., 2025). Yet SROI typically starts from interventions and outcomes, not from micro time allocations within organisational processes. Even recent systematic reviews highlight persistent methodological variation and proxy justification challenges (Sillcox et al., 2026). Therefore, SROI requires a disciplined "input accounting" layer to connect AI-enabled released time to intervention design and evidence.

Existing approaches measure productivity, cost, or social outcomes separately, but do not provide an integrated and auditable mechanism linking these dimensions within organisational decision-making.

3 Released Time (RT): Definition and TDABC-Based Quantification

TDABC provides the technical foundation for quantifying Released Time (RT), extending time equations with complexity drivers and valuing capacity through capacity cost rates (Etges, Polanczyk and Urman, 2020; Etges et al., 2024). We define Released Time in period t as the sum of baseline-to-post reductions in task time, adjusted by a quality-risk-compliance factor that preserves neutrality with respect to service quality and operational risk:

$$RT_t = \sum_{i \in T} \left(T_{i,t}^{\text{baseline}} - T_{i,t}^{\text{post}} \right) \cdot \varphi_{i,t}$$

where T is the set of tasks affected by AI, $T_{i,t}^{\text{baseline}}$ is the standard execution time of task $i$before AI implementation, $T_{i,t}^{\text{post}}$ is the observed execution time after AI implementation, $\varphi_{i,t} \in (0,1]$ is the quality–risk–compliance adjustment factor.

The factor operationalises a core governance premise: Released Time is counted only when time reductions are compatible with quality standards, safety requirements, and compliance controls, a requirement particularly evident in regulated settings where technology can reduce effort while increasing verification obligations (Mátó and Vilmányi, 2024).

Released Time can be treated as a governable stock only when it is traceable to process data and reconcilable with management control systems. Standardisation and reporting discipline in TDABC implementations are therefore not ancillary; they are enabling conditions for the governance indicator we propose (Shakya et al., 2025).

4 Social Time Reinvestment Ratio (STRR): Allocation and Decomposition

Let K={strategy/coordination, WL, VOL, other} be the set of reinvestment categories and let $a_{k,t}$denote the number of hours allocated to category $k \in K$in period t. The allocation must satisfy:

$$\sum_{k \in K} a_{k,t} \leq RT_t$$

The weak inequality is intentional: it allows conservative governance (unallocated RT pending verification) and supports auditability. We define the Social Time Reinvestment Ratio (STRR) as the share of Released Time reinvested into two governance-salient domains: work-life balance and skills-based volunteering.

$$STRR_t = \frac{a_{\mathrm{WL},t} + a_{\mathrm{VOL},t}}{RT_t}$$

with two decompositions:

$$STRR_t^{\mathrm{WL}} = \frac{a_{\mathrm{WL},t}}{RT_t}$$
$$STRR_t^{\mathrm{VOL}} = \frac{a_{\mathrm{VOL},t}}{RT_t}$$

STRR is a micro KPI (0–1) that makes visible the share of Released Time reinvested in sustainability-consistent categories. STRR reframes AI productivity from a cost narrative into a capacity narrative that can be budgeted and governed. This is particularly relevant given evidence that generative AI may reduce time in communication work (e.g., email) and after-hours work (Dillon et al., 2025), making WL reinvestment a measurable governance outcome rather than an aspirational HR claim.

The decomposition reflects different governance logics. Work-life balance interventions and occupational balance are increasingly studied with systematic evidence on effectiveness and variability (Lentner et al., 2025). Skills-based or occupation-related volunteering, by contrast, links workforce capability to external social value creation and has an expanding conceptual and empirical base (Dempsey-Brench and Shantz, 2022; Biermann, Breitsohl and Meijs, 2024).

5 Governance and Allocation Rules (Who Decides, with What Constraints, and on What Basis)

We now specify a minimal governance architecture that makes STRR implementable and assurance-ready. The proposed model is structured around a cross-functional STRR Steering Group with clearly defined roles and decision rights. The Process Owner (PO) is responsible for validating time equations, baseline and post-AI measurements, and the appropriateness of the quality–risk–compliance adjustment factor ϕ_{i,t}. The Controlling or TDABC Lead (CL) reconciles Released Time with TDABC accounts and underlying capacity logic, ensuring traceability. The HR and Learning Lead (HRL) designs eligible work-life balance (WL) and learning or coordination initiatives while ensuring equitable access across organisational units. The Sustainability or Impact Lead (SIL) curates volunteering portfolios and maintains SROI logic models and proxy mappings (Ashton et al., 2024; Nieto et al., 2024). Ethical oversight is ensured through Worker Representation (WRE), which supervises voluntariness, fairness, and grievance channels.

Eligibility rules and evidence requirements ensure auditability. Every allocation must satisfy a predefined catalogue of admissible activities, including work-life balance practices, structured learning and coordination routines, and verified volunteering roles. Allocations must be supported by verifiable evidence, such as attendance and completion records for learning activities, scheduling data for work-life balance, and timesheets or role documentation for volunteering. A non-substitution rule applies: allocated hours must be attributable to genuinely generated Released Time and cannot represent a relabelling of existing budgets. Reconciliation with TDABC accounts constitutes the primary control mechanism (Etges, Polanczyk and Urman, 2020).

The governance framework is further defined by a set of constraints relating to quality, continuity, equity, and voluntariness. The quality constraint is enforced through the adjustment factor ϕ_{i,t} and service-level verification checks (Mátó and Vilmányi, 2024). Continuity constraints ensure minimum staffing and operational coverage, preventing service degradation. Equity constraints require monitoring of allocation distributions across roles and organisational units. Voluntariness constraints ensure that participation in volunteering activities is strictly opt-in; corresponding evidence and participation rules reflect the established importance of experience quality, leadership, and organisational support for volunteer retention and outcomes (Forner et al., 2024; Holtrop et al., 2024).

The governance process follows a structured quarterly decision cycle. Released Time is first measured through TDABC time equations combined with quality adjustments. A conservative Released Time budget is then declared, allowing for the possibility of unallocated Released Time pending verification. This budget is subsequently allocated across portfolios, including work-life balance and volunteering, according to catalogue priorities. Assurance activities follow, involving verification of evidence logs, computation of STRR, and

reconciliation with process accounts. Finally, learning feedback loops are activated to refine both AI deployment priorities and reinvestment design.

From a control perspective, the proposed governance architecture combines diagnostic and boundary systems (Simons, 1995). Diagnostic controls are reflected in the measurement and monitoring of Released Time and STRR, ensuring comparability and accountability across organisational units. Boundary controls are embedded in eligibility rules, voluntariness constraints, and non-substitution principles, preventing the misuse of Released Time and ensuring alignment with organisational values and stakeholder expectations. This dual control structure strengthens the robustness and credibility of the governance mechanism.

6 From STRR to S-STRR: Micro-To-Macro Linkage via SROI

STRR governs the share of Released Time reinvested but does not itself monetise social outcomes. To evidence societal value, the volunteering portfolio is linked to SROI logic models and proxy maps, explicitly drawing on established guidance while acknowledging methodological variation and uncertainty (Ashton et al., 2024; Sillcox et al., 2026).

We define the Societal Social Time Reinvestment Ratio (S-STRR) as:

$$S\text{-}STRR_t = \frac{\widehat{SV}_t}{RT_t}$$

where $\widehat{SV}_t$ denotes the monetised social value generated in period t, estimated through SROI procedures, including stakeholder identification, outcome mapping, proxy valuation, and adjustments for deadweight, attribution, displacement, and drop-off (Ashton et al., 2024; Gutiérrez-Nieto et al., 2025).

A modular adoption strategy ensures that implementation effort remains proportionate by leveraging management control and HR data (Jorgji et al., 2024), while scaling SROI only as governance maturity increases, moving from STRR-only configurations to pilot SROI applications and, ultimately, to fully developed S-STRR implementations.

7 Design and Early-Stage Validation Protocol

7.1 Design method (conceptual engineering + measurement protocol)

The framework follows a design and validation logic appropriate for conceptual and methodological contributions. It is grounded in a conceptual engineering approach combined with a structured measurement protocol. Within this approach, the core constructs, namely Released Time (RT), the Social Time Reinvestment Ratio (STRR), and the Societal STRR (S-STRR), are specified as governance constructs, providing a structured basis for measuring, allocating, and accounting for AI-enabled Released Time within organisational processes.

Operationalisation is achieved through a set of rules governing time equation measurement, the application of the quality–risk–compliance adjustment factor ϕ_{i,t}, the definition of an eligibility catalogue, the enforcement of evidence discipline, and reconciliation procedures linking allocations to TDABC accounts.

Particular attention is devoted to auditability: allocation logs, decision rights, and assurance routines are explicitly defined and anchored to TDABC reporting discipline, in line with recent standardisation efforts (Etges et al., 2024; Shakya et al., 2025).

The framework is further designed to be modularly implementable, allowing for progressive configurations that reduce reporting burden in early stages while supporting scalability and increasing adoption feasibility over time.

7.2 First-stage test (non-causal): reliability + reconciliation

The initial empirical step is explicitly non-causal and is aimed at assessing whether STRR can be measured reliably and reconciled with process-level accounts. Reliability is evaluated through inter-rater agreement in the classification of allocation activities. Independent raters assign activities to categories $k$based on the eligibility catalogue, with target thresholds set at Cohen's κ ≥ 0.70 or, alternatively, agreement levels of at least 80%, consistent with the exploratory analysis (EA) design.

Reconciliation integrity is assessed by comparing allocation logs with TDABC-derived Released Time totals. The acceptable deviation threshold is set at ≤ 5% at the unit or process level, ensuring consistency between measured Released Time and allocated Released Time. Meeting these thresholds constitutes the condition for progressing to subsequent stages, including field applications and the development of full SROI chains. In the initial pilot, we plan to apply the protocol in some knowledge-intensive, public-facing service unit implementing

AI-assisted drafting, using a quarterly observation window and process-level TDABC accounts as the reconciliation backbone.

Table 1. Auditability evidence map for RT, STRR, and S-STRR. The table links each construct to primary data sources, traceable artefacts (audit trail), and assurance tests/thresholds.

Construct / object	Primary data source(s)	Evidence artefacts (audit trail)	Assurance test(s) / threshold
Released Time (RT)	TDABC time equations; baseline vs post-AI task timing; capacity cost rates; quality–risk–compliance factor ϕ	Time-equation specification sheet; parameter change log; baseline/post measurement log; evidence for ϕ (quality checks, compliance controls)	Traceability to process data; sign-off by Process Owner; stability checks on parameters across periods
Allocation categories (k) and eligibility	Eligibility catalogue; policy/rules repository; volunteering role descriptions; WL/learning initiative registry	Versioned catalogue; decision rights and minutes; admissibility checklists; non-substitution declarations	Catalogue completeness review; exception log review; evidence sufficiency sampling
STRR (0–1)	Allocation log of hours $a_{k,t}$; RT totals from TDABC; HR/timekeeping records (where available)	Signed allocation log; supporting evidence (attendance/completions, schedules, volunteering timesheets); computation worksheet	Inter-rater agreement on category assignment: Cohen's κ ≥ 0.70 (or ≥80% agreement); reconciliation to TDABC RT: deviation ≤ 5%
S-STRR (portfolio → SROI)	SROI logic model; outcome indicators; proxy map; adjustment parameters (deadweight, attribution, displacement, drop-off)	Outcome map; proxy justification file; stakeholder register; assumptions register; sensitivity/uncertainty notes	Conservatism checks on assumptions; documentation of proxy sources; reproducibility of valuation steps; governance sign-off by Impact Lead

Note: RT = Released Time; STRR = Social Time Reinvestment Ratio; S-STRR = Societal STRR; κ denotes inter-rater agreement (Cohen's kappa). Reconciliation threshold refers to deviation between allocation logs and TDABC-derived RT totals (≤ 5%).

7.3 Construct validity boundaries

The construct validity of STRR is defined in terms of its intended scope and limits. STRR claims the ability to make visible and measurable quality-adjusted Released Time, to support auditable allocation processes, and to provide governance visibility over the reinvestment of organisational resources. However, it does not claim causal impact on organisational or societal outcomes at this stage. This limitation is consistent with the exploratory scope of the framework and reflects the methodological positioning adopted in this initial phase of development.

8 Illustrative Implementation Pathway (Non-Empirical)

To provide a concrete illustration of the proposed framework, an illustrative and non-empirical implementation pathway is presented. The setting is a public-facing service unit that introduces AI-assisted drafting for citizen requests. Field evidence suggests that such tools can reduce time spent on writing-intensive activities and email, with potential reductions in after-hours work (Dillon et al., 2025).

The implementation begins with process mapping and the quantification of Released Time. The unit identifies and maps key activities, including triage, drafting, verification, and escalation. Baseline and post-AI execution times are measured, and the quality–risk–compliance adjustment factor ϕ_{i,t} is applied to discount any apparent time savings that require additional verification for compliance purposes. The resulting output is the measure of Released Time, RT_t.

On this basis, governance allocation is carried out by the STRR Steering Group. Released Time is distributed across three main categories: work-life balance interventions, such as scheduling buffers aimed at reducing after-hours work; learning and coordination activities, including structured "knowledge clinics" designed to codify recurrent cases; and skills-based volunteering initiatives, for example through partnerships that support digitally excluded citizens.

Following allocation, all activities are documented through verifiable evidence, enabling the computation of STRR and its decompositions. This allows the organisation to compare units, identify where AI adoption effectively releases time, and make reinvestment decisions both visible and budgetable within a governance framework.

In a subsequent stage, the volunteering portfolio may be linked to a simplified SROI model, supported by proxy guidance and conservative adjustments in line with established practice (Ashton et al., 2024; Sillcox et al., 2026). This enables the transition from internal capacity governance to the estimation of monetised social value.

Overall, the scenario illustrates the full pathway connecting measurement and governance, showing how AI adoption leads to Released Time, how allocation rules translate into STRR, and, where implemented, into S-STRR. This integrated sequence constitutes the core contribution of the framework by making explicit the link between measurement, allocation, and governance.

9 Discussion: Contributions and Implications

Theoretical contribution. We extend debates on AI productivity by formalising a distinction between Created Value and Released Time: AI releases time, but social capacity emerges only through governed reinvestment. This is consistent with evidence that AI affects work patterns and task allocation in heterogeneous ways (Brynjolfsson, Li and Raymond, 2025; Hoffmann et al., 2025).

The organisational capacity enabled by Released Time can be interpreted as a form of technology-enabled slack. Classic contributions have shown that slack is not merely inefficiency but can support innovation and adaptability when appropriately managed (Bourgeois, 1981). However, unlike traditional slack, AI-generated released time is process-based, time-specific, and contingent on technological deployment. Its value is therefore not inherent but depends on governance choices. This reframing extends slack theory by introducing a measurable and auditable form of slack that can be intentionally directed towards both organisational and societal objectives.

This argument also resonates with the attention-based view of the firm, according to which organisational behaviour depends on how decision-makers allocate and channel attention (Ocasio, 1997). AI may compress execution time, but it does not determine how the resulting capacity is used. The allocation of activities enabled by Released Time therefore becomes an attention-allocation problem: which activities, stakeholders, and time horizons are prioritised. In this sense, STRR functions not only as a measurement device, but also as a cognitive and governance mechanism that makes the destination of organisational attention explicit and accountable.

Managerial contribution. STRR functions as a governance-ready KPI to steer reinvestment portfolios and prevent the “efficiency trap” in which Released Time is absorbed invisibly. Given evidence that volunteering quality, leadership, and role alignment influence retention and outcomes, governance constraints are not optional but constitute the core mechanism (Forner et al., 2024; Bansi et al., 2025).

Reporting/assurance relevance. STRR is aligned with the need for workforce-related disclosures (ESRS S1 includes metrics and targets for training/skills and work-life balance), and with global human capital reporting standards (ISO 30414:2025).

Finally, the framework is consistent with the concept of absorptive capacity, which emphasises the role of organisational capabilities in recognising, assimilating, and applying new knowledge (Cohen and Levinthal, 1990). By reallocating Released Time towards learning, coordination, and knowledge-sharing activities, organisations can transform AI-enabled efficiency gains into enhanced absorptive capacity, thereby strengthening long-term adaptability and innovation potential.

The framework therefore redefines AI productivity not as a cost phenomenon, but as a governance problem centred on the allocation and accountability of organisational capacity.

10 Limitations and Research Agenda

This paper is deliberately limited to a first-stage, non-causal validation of STRR (reliability + reconciliation). Future research should extend the framework through field studies, causal contrasts, longitudinal evaluation of retention and learning outcomes, and full SROI chains with sensitivity analysis, consistent with the methodological variability highlighted in SROI reviews (Sillcox et al., 2026).

11 Conclusions

Generative AI can free time, but organisations require a governance mechanism to ensure that Released Time becomes sustainable organisational capability and, where relevant, societal value. By quantifying Released Time through TDABC and introducing STRR as an auditable reinvestment ratio, extended to S-STRR through disciplined SROI integration, this paper provides a bridge between measurement and governance that operationalises the shift from cost to social capacity.

12 Ethics Declaration

No ethical clearance was required for this conceptual and methodological paper. (Future field studies will seek relevant approval where required.)

13 AI Declaration

AI tools were used for language refinement and drafting support. All constructs, formulae, governance rules, and methodological choices were defined and critically validated by the authors, who retain full responsibility for the paper.

References

Ashton, K., Cotter-Roberts, A., Clemens, T., Green, L. and Dyakova, M. (2024) 'Advancing the SROI framework to capture the social value of public health interventions', *Public Health*, 226, pp. 122–127

Bansi, D., Petkari, E., Sánchez-Gutiérrez, T. and Pinto da Costa, M. (2025) 'A systematic review of how volunteering relates to job satisfaction and job performance', *Discover Psychology*, 5, 90.

Biermann, K.Y., Breitsohl, H. and Meijs, L.C.P.M. (2024) 'Occupation-Related Volunteering: A qualitative systematic literature review, conceptualization, and directions for future research', *SAGE Open*, 14(2), pp. 1–24.

Bourgeois, L.J. (1981) 'On the measurement of organizational slack', *Academy of Management Review*, 6(1), pp. 29–39.

Brynjolfsson, E., Li, D. and Raymond, L.R. (2025) 'Generative AI at work', *The Quarterly Journal of Economics*, 140(2), pp. 889–942.

Cohen, W.M. and Levinthal, D.A. (1990) 'Absorptive capacity: A new perspective on learning and innovation', *Administrative Science Quarterly*, 35(1), pp. 128–152.

Dempsey-Brench, K. and Shantz, A. (2022) 'Skills-based volunteering: A systematic literature review of the intersection of skills and employee volunteering', *Human Resource Management Review*, 32.

Dillon, E.W., Jaffe, S., Immorlica, N. and Stanton, C.T. (2025) 'Shifting work patterns with generative AI', *NBER Working Paper* 33795.

Edvinsson, L. and Malone, M.S. (1997) *Intellectual capital: Realizing your company's true value by finding its hidden brainpower*. New York: HarperBusiness.

Etges, A.B., Jones, P., Liu, H., Zhang, X. and Haas, D. (2024) 'Improvements in technology and the expanding role of TDABC to increase value in provider organisations: A literature review', *Frontiers in Pharmacology*, 15.

Etges, A.B., Polanczyk, C.A. and Urman, R.D. (2020) 'A standardized framework for TDABC in healthcare: The TDABC in Healthcare Consortium consensus statement', *BMC Health Services Research*, 20, 1107.

European Financial Reporting Advisory Group (EFRAG) (2023) *ESRS S1 Own Workforce* (Delegated Act Annex).

Forner, V.W. et al. (2024) 'Predictors of turnover amongst volunteers: A systematic review and meta-analysis', *Journal of Organizational Behavior*, 45(3), pp. 434–458.

Gutiérrez-Nieto, B., Camón-Cala, J., Cuéllar-Fernández, B. and Fuertes-Callén, Y. (2025) 'A bibliometric analysis of the social return on investment', *Humanities & Social Sciences Communications*.

Hoffmann, M., Boysel, S., Nagle, F., Peng, S. and Xu, K. (2025) *Generative AI and the nature of work* (working paper).

Holtrop, D. et al. (2024) 'Exploring volunteer turnover reasons, intentions, and behavior', *Group & Organization Management*.

International Organization for Standardization (ISO) (2025) *ISO 30414:2025 Human resource management - Requirements and recommendations for human capital reporting and disclosure*.

Jorgji, S., Teta, J., Mousa, S. et al. (2024) 'Sustainable human capital management, ESG, and firm performance: Moderating role of ESG disclosure', *Journal of Human, Earth, and Future*, 5(2).

Lentner, S. et al. (2025) 'Interventions promoting occupational balance in adults: A systematic literature review', *PLOS ONE*, 20(6).

Mátó, Á.R. and Vilmányi, M. (2024) 'Evidence-based experiences of using TDABC in telemedicine-based healthcare delivery protocols', *Telemedicine and e-Health*.

Nieto, I. et al. (2024) 'Striving for global consensus: A systematic review of SROI applied to physical activity and sport', *Journal of Physical Activity and Health*, 21(12), pp. 1246–1262.

Nonaka, I. and Takeuchi, H. (1995) *The knowledge-creating company: How Japanese companies create the dynamics of innovation*. Oxford: Oxford University Press.

Ocasio, W. (1997) 'Towards an attention-based view of the firm', *Strategic Management Journal*, 18(S1), pp. 187–206.

Shakya, S. et al. (2025) 'TDABC and its use in health economic analysis: A systematic literature review', *Applied Health Economics and Health Policy*, 23, pp. 989–1004.

Sillcox, C. et al. (2026) 'Social return on investment of interventions supporting aging in place: A systematic review', *Journal of the American Medical Directors Association*, 27(6), 106177.

Simons, R. (1995) *Levers of control: How managers use innovative control systems to drive strategic renewal*. Boston, MA: Harvard Business School Press.

How European Living Labs Contribute To the Creation of Ecosystems of Innovation: A Social Network Analysis (SNA) of Project-Based Collaboration

Sara Gigli[1], Anna Turchetta[2]
[1] University of Cassino and Southern Lazio, Cassino, Italy
[2] Independent PhD, Aquino, Italy

Abstract

In recent years, increasing attention has been devoted to the role of collaboration and innovation in addressing sustainability challenges. Sustainable innovation has become a key driver of long-term value creation. Technological progress and knowledge exchange are crucial enablers of such innovation, especially within the open innovation framework, where firms integrate internal and external knowledge sources.

Within this evolving landscape, living labs have emerged as pivotal actors in innovation ecosystems. Based on principles of co-creation, experimentation, and user involvement, living labs facilitate the development and testing of innovative solutions in real-life contexts. By bringing together multiple stakeholders, they foster collaborative processes to tackle complex societal challenges and advance sustainability transitions.

This study addresses the following research question: *How do knowledge collaboration networks among living labs contribute to territorial innovation ecosystems?* To answer this question, we adopt a Social Network Analysis (SNA) approach, which is particularly suitable for examining patterns of collaboration and knowledge exchange. Data were manually collected from the websites of living labs listed in the European Network of Living Labs (ENoLL) database. The analysis focuses on four European countries (France, Germany, Italy, and Spain) selected due to their high concentration of living labs and the availability of comparable data.

In the constructed network, nodes represent organizations involved in living lab activities, and edges represent project-based collaborations. Each tie is characterized by attributes such as the number of shared projects, collaboration intensity, domain, and interaction type. Using RStudio, we analyse both global network properties and node-level centrality measures to assess the structural role of living labs within the ecosystem.

Preliminary findings reveal a partially connected network characterized by heterogeneous collaboration patterns. While the overall network density is relatively low, a very small number of living labs act as central hubs. This study contributes to the literature on knowledge management and innovation ecosystems by providing empirical evidence on the relational mechanisms underpinning knowledge-driven innovation. However, limitations related to data availability suggest the need for future research integrating longitudinal and qualitative approaches to better capture the evolution of living lab networks.

Keywords: Living labs; Innovation ecosystems; Knowledge collaboration; Social Network Analysis (SNA)

Paper type: Academic Research Paper

1 Introduction

To achieve the SDGs by 2030, national and international strategies have increasingly emphasized sustainability challenges in business (Zbuchea et al., 2024). Collaboration among different actors in society has had a positive impact on sustainability. For instance, knowledge spillovers generated by industrial collaboration have accelerated the low-carbon transition (Wu and Xu, 2025; Truong et al., 2026). Innovations that improve sustainability performance across social, ecological, and economic dimensions are referred to as sustainable innovations (Boons et al., 2013; Perriton et al., 2025). In this scenario, technology is increasingly essential for driving business innovation (Munteanu et al., 2025).

Schumpeter defined innovation as the motor of the economy (1934). In contemporary societies, innovation and knowledge creation are accelerating (De Jong et al., 2010; Lehmann et al., 2015). Particularly, technological innovation is crucial for sustainable development amid globalization and climate change (Yi et al., 2023; Truong et al., 2026). Innovation is not limited to the output of an internal, closed company process. It is born of ideas, knowledge, and technologies from both within and beyond firm boundaries. This new approach to innovation was first defined by Chesbrough et al. (2005) as open innovation (OI) (Chesbrough, 2006; Schuurman et al., 2016; Sergianni et al., 2024; Pacheco et al., 2025). According to this new framework, innovation is a nonlinear process in which internal R&D departments and the outside world collaborate. In this context, companies benefit from collaboration synergies.

In this scenario, the living lab plays a crucial role in an innovation ecosystem (Tagliazucchi et al., 2024; Najda-Janoszka et al., 2025). Living labs build on the principles of co-creation, exploration, experimentation, and innovation, where user-citizens are at the source of creation (Roy, 2011; Ondiek and Moturi, 2019). They have arisen as "a prominent approach to structure and stimulate innovation by testing new sociotechnical arrangements *in situ* and at a meso-scale" (Engels et al., 2019, p. 1). According to Leminen et al.'s (2012) empirical investigation, there are four leading living lab actors: utilizers, enablers, providers, and users. Through living labs, it is possible to co-create participatory approaches to deal with wicked challenges (Wamsler et al., 2020; Paskaleva and Cooper, 2021; Tiwari et al., 2025). Living labs operate sustainability transformation through experimentation (Cacciolatti et al., 2025).

Our research question is:

"How do knowledge collaboration networks among living labs contribute to territorial innovation ecosystems?"

To carry out our research, we applied Social Network Analysis (SNA) because it is well-suited to analysing collaboration networks (Borgatti and Li, 2009; Arranz et al., 2020). Originally, SNA was used in sociology (Moreno, 1934; Lewin, 1936), but since the 1990s, it has also been introduced in management research (Zheng et al., 2016). Data were collected manually by consulting the websites of living labs. We downloaded the list of living labs, including those in the European Network of Living Labs (ENoLL) database. We consider only France, Germany, Italy, and Spain as these countries show a high concentration of living labs and provide sufficiently comparable and reliable data on project-based collaborations. For each living lab, we analysed projects that they carried out. In our network, nodes are represented by organizations connected by project-based collaborations (edges). Edges represent collaborative ties established through shared projects. They are enriched with multiple attributes, including the number of shared projects, collaboration intensity (low, medium, high), collaboration domain (e.g., agrifood, sustainability, digital innovation), type of collaboration, and tie weight, capturing the strength of inter-organizational relationships. We used RStudio to process the data. Using SNA techniques, the study aims to examine both global network properties and node-level centrality measures to assess the structural role of living labs within the collaboration network.

Preliminary findings reveal a partially connected network characterized by heterogeneous collaboration patterns. While the overall network density is relatively low, a limited number of living labs emerge as central hubs, concentrating a significant share of collaborations, while most organizations remain weakly connected. In addition, cross-regional and cross-sectoral collaborations underscore the role of these key actors in bridging institutional and geographical boundaries. Theoretically, our research contributes to the knowledge management and innovation ecosystem literature by providing empirical evidence on the relational mechanisms through which living labs enable knowledge-driven innovation. In addition, the results underscore the strategic importance of supporting living labs as ecosystem orchestrators, capable of strengthening collaborative intensity. Not all living labs analysed have a website section where all projects and partners can be consulted. These represent key limitations of this research because we can't map all effective collaboration. Finally, future research could extend the analysis longitudinally and integrate qualitative insights better to capture the evolutionary dynamics of living lab-based innovation ecosystems.

2 Literature Review

Innovation, defined by Schumpeter (1934) as the motor of the economy, is accelerating rapidly in our contemporary society (De Jong et al., 2010; Lehmann et al., 2015). At the beginning of the 2000s, Chesbrough et al. (2005) introduced the concept of open innovation (OI). The concept of open innovation highlights practices that cross organizational boundaries to accelerate the creation and appropriation of value (Pacheco et al., 2025). Indeed, OI "enables the pooling of resources and capabilities from various companies and organisations of all kinds" (Sergianni et al., 2024, p. 189).

Living labs are born on the philosophy of the OI paradigm and are considered open innovation networks (Sergianni et al., 2024). They are open innovation ecosystems that support research and innovation activities by embedding them within real-world contexts (Hossain et al., 2019; Tagliazucchi et al., 2024). Nowadays, Europe has a high confluence of operating living labs (Hossain et al., 2019).

In accordance with living labs' philosophy, experimentation is the key mechanism to generate sustainable transformation (Cacciolatti et al., 2025). Living labs assume a pivotal role within innovation ecosystems nowadays (Tagliazucchi et al., 2024; Najda-Janoszka et al., 2025). They are grounded in principles such as co-creation, exploration, experimentation, and innovation, with citizens and users actively involved in generating value (Roy, 2011; Ondiek and Moturi, 2019). Living labs have emerged as "a prominent approach to structure

and stimulate innovation by testing new sociotechnical arrangements *in situ* and at a meso-scale" (Engels et al., 2019, p. 1).

Drawing on the empirical work of Leminen et al. (2012), four main categories of actors can be identified within living labs: utilizers, enablers, providers, and users. These environments enable the development of participatory and collaborative approaches to address complex and "wicked" societal challenges (Wamsler et al., 2020; Paskaleva and Cooper, 2021; Tiwari et al., 2025). Furthermore, living labs contribute to sustainability transitions by fostering experimentation and iterative learning processes (Cacciolatti et al., 2025).

3 Methodology

To answer our research question, we applied Social Network Analysis (SNA) (Moreno, 1934; Lewin, 1936; Borgatti et al., 2009; Zheng et al., 2016).

According to this methodology, organizations are conceptualized as nodes connected through relational ties (Borgatti and Li, 2009; Borgatti et al., 2009; Arranz et al., 2020). In our study, ties are operationalized as project-based collaborations, representing discrete interactions that serve as proxies for underlying relational structures and knowledge exchange processes. To construct the network, we developed a one-mode network of living labs, where ties indicate shared participation in at least one project.

We started our analysis of living labs, focusing on those included in the European Network of Living Labs (ENoLL). We considered only ENoLL members from Italy, France, Germany, and Spain. For each of them, we consulted their websites to find projects they had carried out. When we identified these projects, we began gathering more information about all partners who had participated. For all members, we assigned a unique ID to each node. This was the most difficult passage because not all websites showed detailed information on their project partners. All edges are enriched with multiple attributes, including the number of shared projects, collaboration intensity (low, medium, high), collaboration domain (e.g., agrifood, sustainability, digital innovation), collaboration type, and tie weight, which captures the strength of inter-organizational relationships.

After this first phase of our research, when all data were collected manually, we began the analysis in RStudio. There, we calculated key network metrics, including density and centrality measures. In this process, we applied the SNA protocol (Moreno, 1934; Lewin, 1936; Borgatti et al., 2009; Zheng et al., 2016).

4 Findings and Discussion

The social network analysis (SNA) reveals a network composed of 167 nodes and 163 edges, indicating a relatively sparse structure. The network density is very low (0.0118), suggesting limited collaboration among actors. Furthermore, the global clustering coefficient results equal zero, indicating the absence of cohesive subgroups and triadic closure. This implies that collaborations are predominantly dyadic and not embedded within broader collaborative clusters. Finally, the network diameter is 4, suggesting that, despite low connectivity, the network remains relatively compact (see Table 1).

Table 1**:** Social Network Analysis (SNA) metrics

Metric	Value
Nodes	167
Edges	163
Density	0.0118
Clustering coefficient	0
Diameter	4

Overall, these results indicate that the network cannot be considered a fully developed collaborative ecosystem. Rather, it is characterized by weak connectivity, a limited number of relationships, and the absence of cohesive clusters. This suggests that interactions among living labs are fragmented and lack an integrated structure.

To further explore the network's structure, degree centrality was used to identify the most influential actors. The results highlight a highly uneven distribution of connections, with a small number of nodes serving as central hubs, while most actors exhibit very low connectivity.

Table 2: Top hubs in the living labs network

Node	Country	Degree
Santa Chiara Lab	Italy	38
SSI living lab - Home care lab	Spain	36
VidAAS Centre	Spain	30
Gérontopôle Nouvelle-Aquitaine	France	20
K8 Institute for Strategic Aesthetics	Germany	18

These findings suggest that the network is highly centralized, relying on a limited number of key actors rather than a distributed, collaborative structure. Consequently, interactions are concentrated around these hubs, rather than evenly spread across the network.

The network visual representation (Figure 1) further supports these results. It reveals a fragmented configuration composed of multiple disconnected components, each typically organized around a central node and several peripheral actors. The visualization reveals a hub-and-spoke structure, with several disconnected components organized around central actors, further confirming the high level of network centralization.

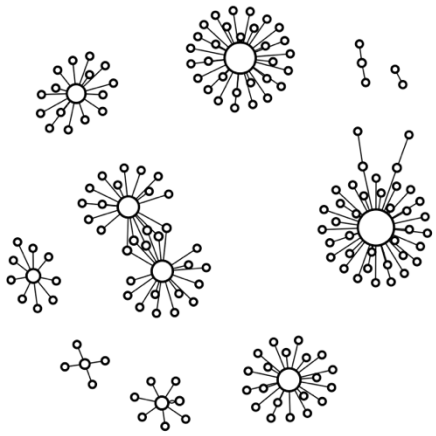


Figure 1: Living Labs Collaboration Network

Living labs are described as platforms where heterogeneous actors collaborate to develop and test innovative solutions (Hossain et al., 2019; Leminen et al., 2016). Within this perspective, they are expected to operate as interconnected ecosystems enabling distributed knowledge flows and collaborative innovation.

However, the empirical evidence emerging from this study provides a more nuanced and, in part, contrasting perspective. The social network analysis reveals a fragmented, weakly connected structure characterized by low density, the absence of cohesive subgroups, and strong centralization around a limited number of actors. These findings suggest that collaboration among living labs is not evenly distributed and does not translate into a fully integrated ecosystem. This result contrasts with several contributions that emphasize the collaborative potential of living labs in addressing complex societal challenges. For example, living labs are often portrayed as enabling stakeholder collaboration and participatory decision-making, particularly in sustainability and climate-related contexts (Tiwari et al., 2025). Similarly, studies on urban agroecosystem living labs highlight their role in fostering co-creation, stakeholder engagement, and sustainable development outcomes (Riaño-Herrera et al., 2025). These perspectives tend to describe living labs as cohesive and effective collaborative environments.

Nevertheless, a growing body of literature suggests that such outcomes are not automatic but depend on specific structural and organizational conditions. Co-creation processes require favourable configurations of cognitive, organizational, and contextual factors, as well as effective mechanisms for integrating knowledge (Centeno et al., 2026). From this perspective, the effectiveness of living labs depends on the presence of enabling conditions. The results of this study align with this view, indicating that the absence of strong and widespread connections may hinder the emergence of integrated collaboration.

Moreover, prior research highlights several structural and managerial challenges associated with living labs. These include governance complexity, scalability issues, stakeholder heterogeneity, and difficulties in sustaining collaboration over time (Hossain et al., 2019; Gibert and Plantec, 2026). In particular, the role of living labs as intermediaries requires them to coordinate actors with divergent interests and expectations, making collaboration inherently complex. These challenges are consistent with the network structure observed in this study. The presence of multiple disconnected components and hub-dominated configurations suggests that

collaboration is localized and does not easily scale across the broader ecosystem. In line with this, previous studies on urban experimentation and living labs also question their transformative capacity, highlighting that their impact is often difficult to measure and may be limited by institutional and structural constraints (Kropp, 2025). Similarly, research on circular economy living labs identifies coordination challenges, infrastructural limitations, and difficulties in aligning stakeholders as key barriers to effective collaboration (Elbroumi et al., 2025). These insights help explain why, despite their theoretical promise, living labs may fail to generate fully integrated innovation ecosystems in practice.

5 Conclusions

Living labs are increasingly recognized as key actors within innovation ecosystems. They facilitate fostering collaboration, co-creation, and knowledge exchange among diverse stakeholders (Roy, 2011; Leminen et al., 2012; Wamsler et al., 2020). In line with the literature on open and sustainable innovation, they are expected to facilitate distributed and interconnected knowledge dynamics that contribute to territorial innovation ecosystems (Chesbrough, 2006; Boons et al., 2013; Pacheco et al., 2025).

However, our findings provide a more complex perspective. The empirical evidence from the social network analysis reveals that, rather than forming a cohesive and integrated ecosystem, living labs are embedded in a fragmented and weakly connected network. Collaboration patterns appear limited and unevenly distributed, with interactions concentrated among a small number of central actors, while most living labs remain only marginally connected. These results suggest a discrepancy between the theoretical expectations of living labs as collaborative innovation ecosystems and their actual structural configuration. Instead of a distributed, interconnected network, the observed structure is characterized by isolated clusters and hub-dominated interactions, suggesting that knowledge-exchange processes may be less pervasive and inclusive than expected.

This study aims to contribute to the literature on knowledge management and innovation ecosystems by highlighting the importance of network structure in shaping collaborative dynamics (Borgatti and Li, 2009; Arranz et al., 2020). Specifically, it shows that the mere presence of collaborative initiatives does not necessarily translate into an integrated ecosystem; rather, it may result in fragmented configurations with limited systemic impact. Our findings underscore the need for policymakers and ecosystem orchestrators to foster stronger, more widespread connections among living labs. Enhancing interconnectivity and promoting cross-cluster collaboration could improve knowledge diffusion, increase innovation capacity, and better support the achievement of sustainability goals (Zbuchea et al., 2024; Wu and Xu, 2025).

This study is subject to some limitations. Data were manually collected from publicly available sources, and not all living labs provide complete information on their projects and partners. As a result, some collaborative ties may not have been captured. Future research could extend this work by adopting longitudinal approaches and integrating qualitative methods to understand better the evolution and governance of living lab networks over time.

6 Ethics declaration

This study did not require ethical approval because we used public data.

7 AI declaration

No AI-assisted technologies were used in this manuscript.

References

Arranz, N., Arroyabe, M.F. and Schumann, M. (2020) "The role of NPOs and international actors in the national innovation system: A network-based approach", *Technological Forecasting and Social Change*, 159, 120183.

Boons, F., Montalvo, C., Quist, J. and Wagner, M. (2013) "Sustainable innovation, business models and economic performance: an overview", *Journal of Cleaner Production*, 45, pp 1-8.

Borgatti, S.P. and Li, X. (2009) "On social network analysis in a supply chain context", Journal *of supply chain management*, Vol. 45, No. 2, pp 5-22.

Borgatti, S.P., Mehra, A., Brass, D. and Labianca., G. (2009) "Network Analysis in the Social Sciences", *Science*, 323, pp. 892-895.

Cacciolatti, L., Lee, S.H., Christodoulou, I. and Christofi, M. (2025) "Living labs as ethical spaces: Fostering innovation and sustainability in food systems", *Journal of Business Ethics*, pp 1-21.

Centeno, J.P., Steen, T., and Müller, A.P.R. (2026) "Co-creation pathways to open-social innovation: Assessing knowledge integration, organisational settings and contextual conditions in European Living Labs", *Technological Forecasting and Social Change*, 228, 124705.

Chesbrough, H. (2005) "Open innovation: a new paradigm for understanding industrial innovation", in Chesbrough, H., Vanhaverbeke, W. and West, J. (Eds.), Open innovation: researching a new paradigm, Oxford University Press, Oxford.

Chesbrough, H. (2006) *Open business model: How to thrive in the new innovation landscape*, Harvard Business Press.

De Jong, J.P., Kalvet, T. and Vanhaverbeke, W. (2010) "Exploring a theoretical framework to structure the public policy implications of open innovation", *Technology Analysis & Strategic Management*, Vol. 22, No. 8, pp 877-896.

Elbroumi, S., Idrissi, M.A., Eddahmouny, H., and Chaaouan, M. (2025) "Living labs supporting circular cities in Morocco: Towards collaborative waste management in Casablanca", *Sustainable Futures*, 10, 100868.

Engels, F., Wentland, A. and Pfotenhauer, S.M. (2019) "Testing future societies? Developing a framework for test beds and living labs as instruments of innovation governance", *Research Policy*, Vol. 48, No. 9, 103826.

Gibert, C., and Plantec, Q. (2026) "The living lab and the cursed catalyst: Navigating the legitimacy challenges of innovation intermediaries for sustainable innovation", *Research Policy*, Vol. 55, No. 3, 105403.

Hossain, M., Leminem, S., and Westerlund, M. (2019) "A systematic review of living lab literature", *Journal of Cleaner Production*, 213, pp. 976–988.

Kropp, C. (2025) "Modernist social imaginaries in research funding for urban experimentation–a qualitative analysis of expectations and visions in German-language funding calls", *Research Policy*, Vol. 54, No. 9, 105306.

Lehmann, V., Frangioni, M. and Dubé, P. (2015) "Living Lab as knowledge system: an actual approach for managing urban service projects?", *Journal of Knowledge Management*, Vol. 19, No. 5, pp 1087-1107.

Leminen, S., Nyström, A.G., Westerlund, M., and Kortelainen, M.J. (2016) "The effect of network structure on radical innovation in living labs", *Journal of Business & Industrial Marketing*, Vol. 31, No. 6, pp. 743-757.

Leminen, S., Westerlund, M. and Nyström, A. (2012) "Living Labs as open-innovation networks", *Technology Innovation Management Review*, Vol. 2, September, pp 6-11.

Lewin, K. (1936) *Principles of Topological Psychology*, McGraw-Hill, New York.

Moreno, J.L. (1934) *Who Shall Survive? Foundations of Sociometry, Group Psychotherapy, and Sociodrama*, Nervous and Mental Disease Publishing Co., Washington, D.C.

Munteanu, I., Ionescu-Feleagă, L., Ionescu, B.Ș., Condrea, E. and Romanelli, M. (2025) "From Metrics to Meaning: Research Trends and AHP-Driven Insights into Financial Performance in Sustainability Transitions", *Sustainability*, 17, 6437.

Najda-Janoszka, M., Kajzer-Bonk, J., Milewska, E. and Wrona, S. (2025) "Integrating science, technology, and experimental knowledge for sustainable innovation: A Living Lab approach to urban biodiversity management", *Journal of Entrepreneurship, Management and Innovation*, Vol. 21, No. 2, pp 33-55.

Ondiek, M.A. and Moturi, C. (2019) "An assessment of the sustainability of Living Labs in Kenya", *Innovation & Management Review*, Vol. 16, No. 4, pp 391-403.

Pacheco, J.N., Turro, A. and Urbano, D. (2025) "Open social innovation: A systematic literature review and future research agenda", *Technological Forecasting and Social Change*, 216, 124160.

Paskaleva, K. and Cooper, I. (2021) "Are living labs effective? Exploring the evidence", *Technovation*, 106, 102311.

Perriton, E., Korsgaard, S., Gustafsson, S., Evald, M.R. and Ingstrup, M.B. (2025) "Not Just Cocreation: Exploring the Shifting Interaction Logics in Sustainable Innovation", *Business Strategy and the Environment*, Vol. 34, No. 4, pp 5095-5113.

Riaño-Herrera, D.A., Romero-Perdomo, F., Quintero-Castro, J., García-Sánchez, A., and Rodríguez-Urrego, L. (2025) "Exploring the design and contributions of urban agroecosystem living labs for sustainable city development", *Sustainable Futures*, 10, 100976.

Roy, J. (2011) "Report on the Workshop on a New Citizen-Centric Information System", Société de Recherche Sociale Appliqué, Québec.

Schumpeter, J.A. (1934) *The theory of economic development: An inquiry into profits, capital, credit, interest, and the business cycle*, Harvard University Press.

Schuurman, D., Baccarne, B., Marez, L.D., Veeckman, C. and Ballon, P. (2016) "Living Labs as open innovation systems for knowledge exchange: solutions for sustainable innovation development", *International Journal of Business Innovation and Research*, Vol. 10, No. 2-3, pp 322-340.

Sergianni, L., De Chiara, A. and Mauro, S. (2024) "Open Innovation as Fuel for the Circular Economy: an Analysis of the Italian Context", *Journal of Innovation Management*, Vol. 12, No. 1, pp 188-204.

Tagliazucchi, G., Della Santa, S. and Gherardini, F. (2024) "Design of a living lab for autonomous driving: an investigation under the lens of the triple helix model", *The Journal of Technology Transfer*, Vol. 49, No. 3, pp 876-899.

Tiwari, A., Rodrigues, L.C., Caruso, R., Ensenado, E.M., Makousiari, E. and Gharbia, S. (2025) "Multi-criteria analysis of ecosystem-based adaptation strategies for flood risk reduction in Irish, Italian, and Turkish Coastal City Living Labs", *Sustainable Futures*, 10, 101506.

Truong, Q.N., Nguyen, P.V. and Hansen, U.E. (2026) "Global research trends in renewable energy technology transfer: implications for innovation policy and emerging economies", *EuroMed Journal of Business*, Vol. 21, No. 5, pp 1-17.

Wamsler, C., Wickenberg, B., Hanson, H., Olsson, J.A., Stålhammar, S., Björn, H., ... and Zelmerlow, F. (2020) "Environmental and climate policy integration: Targeted strategies for overcoming barriers to nature-based solutions and climate change adaptation", *Journal of Cleaner Production*, 247, 119154.

Wu, B. and Xu, Y. (2025) "Towards sustainable development: can industrial collaborative agglomeration become an accelerator for green and low-carbon transformation of resource-based cities in China?", *Journal of Environmental Management*, 381, 125199.

Yi, S., Raghutla, C., Chittedi, K.R. and Fareed, Z. (2023) "How economic policy uncertainty and financial development contribute to renewable energy consumption? The importance of economic globalization", *Renewable Energy*, Vol. 202, pp 1357-1367.

Zbuchea, A., Barna, C., Stănescu, S. and Pînzaru, F. (2024) Exploring the Motivations Behind Circular Social Entrepreneurs in Romania. In *Entrepreneurial Motivations: Strategies, Opportunities and Decisions* (pp. 191-203). Cham: Springer Nature Switzerland.

Zheng, X., Le, Y., Chan, A.P., Hu, Y. and Li, Y. (2016) "Review of the application of social network analysis (SNA) in construction project management research", *International journal of project management*, Vol. 34, No. 7, pp 1214-1225.

Legality Rating and Performance of Italian Companies: Face or Mask?

Fabrizio Rossi[1], Paola Rossi[1], Michele Bertoni[1], Domenico Celenza[2]
[1] University of Trieste, Trieste, Italy
[2] University of Cassino and Southern Lazio, Cassino, Italy

Abstract

With Decree-Law of January 24, 2012, Italy introduced the Legality Rating to encourage small and medium-sized enterprises to adopt ethical behaviours and comply with principles of legality. Awarded by the Italian Competition Authority (AGCM) on a scale from one to three stars, the Rating provides certified firms with several advantages, including preferential conditions in accessing bank credit, reputational improvements, and access to public funding. Within a civil-law jurisdiction characterised by concentrated ownership and weak formal investor protection, the Rating has been theoretically conceived as a credible signalling and governance device that should reduce information asymmetries, constrain managerial discretion, and align corporate conduct with ethical standards.

Prior empirical evidence on the Legality Rating remains mixed: some studies document better profitability, lower insolvency risk, and improved earnings quality among certified firms, while others suggest that weaker or less profitable firms may use the Rating as a symbolic management tool. Drawing on agency theory, stakeholder theory, and signalling theory, this study revisits this debate by jointly examining default risk, operational risk, and agency costs.

Using a sample of 1,925 Italian firms holding the Legality Rating, for a total of 7,700 firm-year observations over the period 2020–2023, we apply a fixed-effects econometric model to test the relationship between the Rating score and firm performance. Performance is measured through Altman's Z_SCORE (default risk), the standard deviation of ROA over five years (operational risk), and the ratio of operating expenses to total assets (agency costs), while controlling for size, age, ownership concentration, profitability, leverage, industry, and year effects. We investigate whether firms holding the Rating, and thus considered more virtuous, actually mitigate agency costs and reduce financial distress risk, or whether they hide behind it to pursue window-dressing strategies.

Contrary to expectations, the results do not support the formulated hypotheses. Higher Rating scores are systematically associated with higher operating expenses, greater ROA volatility, and lower Z_SCORE values, suggesting that more highly rated firms exhibit higher operational risk, higher agency costs, and greater proximity to the financial distress zone. These findings raise concerns about the effectiveness of the Rating as a governance and risk-mitigating mechanism and are consistent with a window-dressing interpretation, whereby financially fragile firms strategically use the certification to enhance reputation, attract preferential credit conditions, and conceal underlying weaknesses. The study contributes to the literature on ethical certifications, signalling, and corporate transparency, and offers practical insights for investors, creditors, and minority shareholders.

Keywords: legality rating, Italian firms, firm performance, window dressing, risk-taking

Paper type: Academic Research Paper

1 Introduction

A vast literature has empirically demonstrated that the development of an economic system, firm growth, capital market development, and a country's legal system are closely interrelated (e.g., La Porta et al., 1997). One of the main differences between civil law and common law countries lies in the level of stakeholder protection, which appears higher in common law systems. Firms more virtuous in terms of legality, transparency, and corporate social responsibility are expected to reduce agency conflicts (Jensen and Meckling, 1976) and increase firm value. With Decree-Law of January 24, 2012, Italy introduced the Legality Rating to encourage SMEs to adopt ethical behaviours, providing economic, reputational, and financing benefits to certified firms.

Several studies have investigated the relationship between the Rating and firm performance with mixed results. Caputo and Pizzi (2019) find that the Rating is associated with significantly higher EBITDA and liquidity ratios and lower debt-to-equity ratios. Alfano et al. (2024) document higher profitability and liquidity among certified firms, attributing the differences to reputational effects, and Lo Conte and Sancetta (2023, 2024) confirm these findings. By contrast, La Rosa et al. (2021) show that firms with lower profitability or higher insolvency risk tend to achieve higher rating scores, suggesting that weaker firms use the Rating as a symbolic management tool. They also argue that the Rating is mainly adopted by firms that already comply with legal and ethical standards, and is therefore unlikely to change corporate behaviour.

A growing body of research examines performance through the lens of financial stability. Pizzi et al. (2020), using Altman-type indicators, document that ethically and legally compliant firms exhibit lower insolvency risk. Bertacchini et al. (2025) extend this stream by linking the Rating to lower earnings management, although they note that improved earnings quality may also reflect more conservative accounting practices.

Using a sample of 1,925 Italian firms holding the Legality Rating (7,700 firm-year observations), this study investigates whether rated firms actually mitigate agency costs and reduce financial distress, or whether they hide behind the Rating to pursue window-dressing strategies.

2 Literature Review and Hypothesis Development

2.1 Agency Costs, Corporate Social Responsibility, and the Legality Rating

2.1.1 Agency theory and the ownership–control nexus

The separation of ownership and control is a foundational problem in corporate finance (Berle and Means, 1932; Jensen and Meckling, 1976). Within Jensen and Meckling's (1976) framework, agency costs arise from divergent incentives of principals and agents and comprise monitoring, bonding, and residual loss; they vanish when the owner-manager retains the entire stake and intensify as outside equity increases.

The original Agency Problem Type I focuses on the conflict between dispersed shareholders and professional managers. With concentrated ownership, typical of European economies, the dominant concern shifts to expropriation of minority shareholders by controlling blockholders,Agency Problem Type II (Villalonga and Amit, 2006). La Porta et al. (1999, 2000) document that, outside the United States, most listed firms feature a single controlling shareholder; when that blockholder is an individual or a family, the incentive to extract private benefits makes the Type II problem more relevant. Villalonga et al. (2015) extend this taxonomy in family firms, identifying four agency problems: owner–manager (Type I), controlling–minority shareholder (Type II), shareholder–creditor (Type III), and intra-family (Type IV).

2.1.2 Legal systems, investor protection, and capital markets

The legal environment plays a crucial role in shaping agency dynamics. La Porta et al. (1997) demonstrate a systematic relationship between legal origin and capital market development across 49 countries: common law countries offer stronger investor protection and have larger, more developed markets. Djankov et al. (2008) refine this with the anti-self-dealing index, confirming that countries with weaker provisions exhibit less developed capital markets.

Italy belongs to the civil law tradition and is characterised by relatively weak formal investor protection. Faccio and Lang (2002) document that the majority of continental European firms are controlled by families or the State, with widespread use of pyramidal structures and dual-class shares, a pattern especially pronounced in Italy. The Type II agency problem is therefore particularly relevant, and voluntary mechanisms credibly signalling commitment to legality become especially valuable as complements to thin formal protections.

2.1.3 CSR, stakeholder governance, and agency cost mitigation

A growing literature links CSR to agency cost mitigation: CSR practices reduce information asymmetries, constrain managerial opportunism, and strengthen monitoring (e.g., Ferrell et al., 2016; Jo and Harjoto, 2011). From a stakeholder perspective (Freeman, 1984), CSR aligns managerial interests with those of a broader set of stakeholders through reputational and governance channels.

However, the relationship is not unambiguously positive. Barnea and Rubin (2010) show that insiders with low ownership stakes promote CSR over-investment to enhance personal reputation at the expense of shareholder value. Cespa and Cestone (2007) develop a model in which managers use CSR as an entrenchment mechanism, building stakeholder alliances to insulate themselves from disciplinary takeovers. The agency implications of CSR thus depend on whether ethical practices are adopted as genuine governance tools or as instruments of managerial self-interest.

2.1.4 The Legality Rating as a governance and signalling mechanism

Within the agency-theoretic framework, the Legality Rating operates through two channels: it imposes governance standards that promote transparency and constrain managerial discretion, and it functions as a credible signal to external stakeholders. Signalling theory (Spence, 1973) holds that informed parties can convey unobservable qualities through costly and observable actions; Connelly et al. (2011) emphasise that effective signals must be costly to obtain and difficult to imitate. The Rating satisfies these conditions: it requires verifiable

compliance with legal standards, is awarded by an independent authority (AGCM), and is publicly observable, lowering the cost of external financing.

2.2 Legality Rating and risk

The relationship between CSR and firm risk is a well-established area of inquiry. Building on stakeholder theory (Freeman, 1984), substantial research challenges the Friedmanite view that ethical conduct and economic performance are inherently in tension, arguing instead that they are mutually reinforcing (Porter and Kramer, 2002). CSR benefits encompass improved financial performance, enhanced reputation, lower transaction costs, and reduced firm risk (Weber, 2008). The Legality Rating, by requiring compliance with legal standards verified by an independent authority, can be interpreted as a proxy for CSR in the Italian setting. As Carroll (1991) observed, compliance with the law is a foundational pillar of socially responsible conduct, while McWilliams and Siegel (2001) emphasise that genuine CSR goes beyond mere legal obedience.

2.2.1 CSR, financial stability, and default risk

Empirical evidence consistently supports a negative association between CSR engagement and firm risk. Tsoutsoura (2004) documents that S&P 500 CSR practices are associated with higher accounting performance (ROA, ROE, ROS). Orazalin et al. (2019), focusing on the Russian oil and gas industry, find that sustainability disclosure quality is positively related to Z-scores, indicating greater financial stability among more transparent firms.

A complementary strand examines the link between CSR and default risk across institutional environments. Shahrour et al. (2021), analysing 412 Eurozone firms over 2003–2017, document a negative relationship that intensifies during downturns, providing insurance-like protection. Sun and Cui (2014) reach analogous conclusions, showing that CSR-engaged firms benefit from a lower cost of debt and improved credit ratings. Do (2022), in an international study covering 36 countries, finds that the risk-mitigating effect of CSR is more pronounced for firms with sustained, long-standing commitment.

2.2.2 CSR and risk-taking behaviour

Beyond default risk, an extensive literature investigates the impact of CSR on managerial risk-taking. Orlitzky and Benjamin (2001), in a meta-analytic review, document a robust negative correlation between corporate social performance and financial risk, attributable to trust-building among stakeholders, accumulated social capital, and reduced uncertainty. Mulia and Joni (2020), examining Indonesian listed firms with both operational and market risk measures, find that CSR adoption significantly reduces both, particularly in high-risk emerging-market settings. Harjoto and Laksmana (2018) show that CSR engagement constrains both excessive and insufficient risk-taking, aligning corporate decisions with optimal risk levels.

Taken together, these contributions suggest that CSR operates as a risk-mitigating mechanism by lowering information asymmetries, strengthening trust-based stakeholder relationships, and curbing managerial discretion in risky strategies.

2.2.3 Evidence on the Legality Rating and firm risk

The empirical literature on the Legality Rating, although limited, supports the risk-mitigating interpretation. Caputo and Pizzi (2019) document that voluntary adoption is associated with improved economic, financial, and stock-market performance. La Rosa et al. (2021) argue that the Rating is predominantly adopted by firms whose governance is already attentive to ethical and legal matters. La Rosa (2020) notes that, while in the short term stakeholders reward firms primarily for meeting minimum legality standards, the highest scores serve a longer-term reputational purpose.

Several contributions examine the link between the Rating and firm risk. De Benedetto et al. (2024), using conditional nonparametric frontier analysis, show that the Rating enhances corporate efficiency by mitigating risks, facilitating bank credit access. Lo Conte and Sancetta (2023), comparing 100 service-sector rated firms with a control group, find significant reductions in insolvency risk (higher Z-scores). Pizzi et al. (2020), focusing on 148 firms with the maximum three-star score, reach analogous conclusions and suggest that the Rating may moderate the bankruptcy risk associated with firm age.

2.3 Hypothesis development

2.3.1 Legality Rating and risk

The reviewed evidence is broadly consistent: CSR engagement is systematically associated with lower default risk, lower operational and market risk, and more disciplined risk-taking. The Legality Rating, conceived as a credible signal of legal compliance and as a governance mechanism, can be expected to operate through the same channels (Spence, 1973; Connelly et al., 2011). Accordingly, we formulate:

H1: There is a negative relationship between the Legality Rating score and firms' risk-taking.

2.3.2 Legality Rating and agency costs

Two competing perspectives map onto the "face or mask" question. Under the dominant view, building on agency, stakeholder, and signalling theory, if the Rating represents a genuine commitment to governance quality (the "face" of corporate virtue), then higher scores should correspond to lower agency costs through governance standards constraining managerial discretion, transparency requirements lowering information asymmetries, and the credibility of an independently verified, publicly observable signal (Spence, 1973; Connelly et al., 2011). In concentrated ownership settings with weak formal investor protection (La Porta et al., 1997; Djankov et al., 2008), credible certification may further constrain expropriation.

A contrasting strand treats CSR and ethical certifications as a potential source of agency conflicts. Barnea and Rubin (2010) show managerial over-investment in CSR for personal reputation; Cespa and Cestone (2007) argue that CSR can serve as entrenchment. Applied to the Rating, this suggests certification may function as a "mask" projecting ethical virtue while underlying agency problems persist, with firms with higher agency costs seeking the Rating to compensate for governance weaknesses (La Rosa et al., 2021). We consider this perspective less compelling than in the broader CSR setting, given the Rating's mandatory eligibility requirements and independent verification. We therefore formulate:

H2: There is a negative relationship between the Legality Rating score and agency costs.

3 Methodology and Data Analysis

To test the hypotheses, we constructed a sample of firms holding the Legality Rating, obtained from the Italian Competition Authority (AGCM). Since historical data are not available, the analysis focuses on firms that renewed the Rating between 2022 and 2024 (latest update October 24, 2024). The lack of information on past renewals led us to assume a single renewal per firm.

The sample includes only joint-stock and limited liability companies, excluding partnerships, cooperatives, financial firms, and utilities. These criteria reduced the initial sample of over 10,000 firms to 1,925. The analysis covers 2020–2023, with firm-level financial data from the AIDA database (Bureau van Dijk). All variables were winsorized at the 5th and 95th percentiles.

Firm risk, used to test H1, was measured through Z_SCORE and the standard deviation of ROA. The Z_SCORE has been widely adopted in the legality rating literature (Pizzi et al., 2020; Lo Conte and Sancetta, 2023) and was computed following Altman (1993) for non-listed firms and Altman (1968) for the seven listed firms in the sample. The standard deviation of ROA over five years captures operational risk (Mulia and Joni, 2020; Harjoto and Laksmana, 2018). To test H2, we employed operating expenses (OPEX), as the ratio of operating expenses to total assets, following Ang et al. (2000) as a proxy for agency costs.

The independent variable is the legality rating score, consistent with La Rosa (2020). Star ratings were converted into a numerical scale from 1 to 7 (one star = 1; three stars = 7). Table 1 illustrates the conversion and shows a higher concentration of firms with lower scores.

Table 1: Legality score

Score	Rating	% of firms
Star	1	0.2 %
Star and one "+"	2	5.3 %
Star and two "+"	3	28.4 %
Two stars	4	21.8 %
Two stars and one "+"	5	12.8 %
Two stars and two "+"	6	11.7 %
Three stars	7	19.8 %

Source: Author's own work

Control variables include firm size (log of total assets), age (log of years since establishment), the ownership share of the largest shareholder, profitability (ROA), leverage, and Industry and Year dummies.

Table 2: Description and measurement of variables

Variables	Description	Data Source
Dependent Variables		
Z_SCORE	It was measured according to Altman (1968)	AIDA Bureau van Dijk.
OPEX	Ratio of operating expenses to total assets	AIDA Bureau van Dijk.
SD(ROA)	Standard deviation of ROA over a five-year time horizon	AIDA Bureau van Dijk.
Independent Variables		
RATING	Score varies from 1, if it does not satisfy any requirement of Article 3, to 7, if it satisfies at least 6 requirements of the same article of the conformity provided by AGCM	The Italian Competition Authority (*Autorità Garante della Concorrenza e del Mercato* – AGCM)
Control Variables		
SIZE	Total assets (in logarithmic form)	AIDA Bureau van Dijk.
AGE	Years since firm was established (in logarithmic form)	AIDA Bureau van Dijk.
OWNERSHIP	The percentage of shares held by controlling shareholder	AIDA Bureau van Dijk.
ROA	Yearly EBIT/total assets*100	AIDA Bureau van Dijk.
LEVERAGE	Financial debt divided by total assets	AIDA Bureau van Dijk.
INDUSTRY	Four-digit SIC (Standard Industrial Classification) codes. A dummy variable that takes the value 1 if it belongs to a specific industry sector and 0 otherwise	AIDA Bureau van Dijk.
YEAR	Dummy variables (1, 0) associated with the years	AIDA Bureau van Dijk.

Source: Author's own work

Table 3: Descriptive statistics

Variables	Average	Median	Standard Dev.	Min.	Max.
Z_SCORE	1,87	1,75	0,8	0,59	3,59
OPEX	0,55	0,46	0,33	0,15	1,4
SD(ROA)	3,37	2,55	2,62	0,5	10,18
RATING	4,56	4	1,6	1	7
SIZE	9,25	9,19	1,63	0	16,05
AGE	3,15	3,26	0,67	0	4,66
OWNERSHIP (%)	64,9	60,8	30,1	0,00	100,00
ROA (%)	7,01	5,23	6,25	-0,85	22,78
LEVERAGE	0,57	0,58	0,2	0,19	0,87

Source: Author's own work

In order to verify the effectiveness of the Legality Rating in reducing risk-taking and agency costs, we adopted a fixed-effects econometric model whose equation is as follows:

$$Y_{i,t} = \alpha_i + \beta_1 RATING + \beta_2 SIZE + \beta_3 AGE + \beta_4 ROA_{i,t} + \beta_5 OWNERSHIP + \beta_6 LEVERAGE + \gamma_s + \delta_t + \varepsilon_{i,t}$$

Y_(i,t) represents the dependent variable (Z_SCORE, OPEX, SD(ROA). The independent variable is the Legality Rating (Rating), and ranges from "1" to "7". The control variables are firm size (SIZE), firm age (AGE), percentage of majority shareholder (OWNERSHIP), return on assets (ROA), and firm debt level (LEVERAGE). γ_s represents the sector dummies (Sector), δ_t are the time dummies (Year), and ε_(i,t) is the error term.

4 Results

Table 4 reports the main results. Column (1) shows the regression with Z_SCORE as dependent variable. Since a higher Z_SCORE corresponds to lower default risk, the negative and significant coefficient on Rating (-0.0192, $p < 0.001$) implies that higher rating scores are associated with greater default risk. Column (3) reports results for SD(ROA): the positive and significant coefficient (0.1074, $p < 0.001$) indicates that higher rating scores are associated with greater ROA volatility, suggesting higher operational risk. These findings do not support H1, raising doubts about the effectiveness of the Rating as a risk-mitigating mechanism.

Table 4: Main findings

	Z_SCORE (1)	OPEX (2)	SD(ROA) (3)
Const	2,2343 (0,0736) ***	0,8434 (0,0393) ***	5,9281 (0,3088) ***
RATING	-0,0192 (0,0060) ***	0,0121 (0,0035) ***	0,1074 (0,0269) ***
SIZE	0.0199 (0,0066) ***	-0,0315 (0,0033) ***	-0,1835 (0,0242) ***
AGE	-0,0733 (0,0119) ***	-0,0214 (0,0064) ***	-0,4685 (0,0525) ***
OWNERSHIP	0,0005 (0,0002) **	-0,00005 (0,0001)	0,0023 (0,0010) **
ROA	0,0722 (0,0012) ***	0,0101 (0,0006) ***	0,1627 (0,0058) ***
LEVERAGE	-1,2779 (0,0461) ***	0,4222 (0,0211) ***	-1,4929 (0,1862) ***
INDUSTRY	Yes	Yes	Yes
YEAR	Yes	Yes	Yes
No. Obs.	7700	7700	7700
R-squared	0,684	0,482	0,414
F-test	652,558	141,789	109,767
P-value	[0,000] ***	[0,000] ***	[0,000] ***

Note: The standard error is reported in parentheses. (***) p<0.001, (**) p<0.01, and (*) p<0.05

Source: Author's own work

Column (3) presents the results for the variable SD(ROA), i.e., the standard deviation of ROA, used to measure the firm's operational risk. The coefficient of the Rating variable is positive (0.1704) and statistically significant ($p < 0.001$), suggesting that higher rating scores are associated with greater volatility in ROA. This result indicates that firms meeting more additional requirements exhibit higher operational risk compared to those with a lower legality rating score.

Overall, these findings do not support hypothesis H1, as a higher rating score does not correspond to a reduction in risk-taking, contrary to what would reasonably be expected. The results raise doubts about the effectiveness of the legality rating in reducing firm risk, as firms with higher ratings were expected to be more prudent and less risky.

Column (2) reports the regression results for the dependent variable OPEX, calculated as the ratio of operating expenses to total assets and used as a proxy for agency costs. The coefficient of the Rating variable is positive (0.0121) and statistically significant ($p < 0.001$), indicating that an increase in the rating score is associated with higher operating expenses and, consequently, higher agency costs. H2 is therefore also rejected.

Regarding the control variables, all coefficients are statistically significant ($p < 0.001$). In particular, Size has a positive effect on all three dependent variables, showing a positive coefficient with respect to Z_SCORE and SD(ROA), and a negative coefficient with respect to OPEX. This suggests that larger firms benefit from lower default and operational risk and are more efficient in reducing agency costs.

The ROA variable consistently shows a positive coefficient, indicating that higher profitability improves the Z_SCORE while increasing the volatility of returns on assets.

Conversely, Age shows a negative coefficient across all models, suggesting that older firms experience lower operational risk and reduced operating expenses, but greater exposure to default risk.

Finally, LEVERAGE exhibits a negative coefficient with respect to Z_SCORE and SD(ROA), and a positive coefficient with respect to OPEX. This implies that as leverage increases, firms improve their Z'-score, but agency costs and ROA volatility also increase. Regarding the OWNERSHIP variable, the sign of the coefficient appears ambiguous. On the one hand, a high concentration of ownership would improve the Z_SCORE, but on the other, it would increase risk. This apparent contradiction could be explained by the fact that the controlling shareholder could venture into riskier investments but, at the same time, would ensure prudent financial management that does not jeopardize the company's default.

The results contradict the hypothesised negative relationships. The average Z_SCORE among rated firms is 1.86, comparable to the 1.43 reported by Caputo et al. (2020). These values place rated firms in the Altman "grey area", close to the distress zone (threshold of 1.23).

Two explanations can be proposed. First, beyond economic benefits, firms may use the Rating to mask their financial condition, leveraging the reputational advantage to gain stakeholder trust while concealing fragility.

Firms in serious difficulty may also seek the Rating for easier access to public and bank financing. Similar conclusions are reached by La Rosa et al. (2019), who argue that firms with lower profitability are more likely to pursue higher legality scores to signal a positive attribute that improves the perception of their condition.

These findings raise doubts about whether the Rating is pursued as a genuine certification of transparency or as a form of "window dressing". Maurutto (2024) defines window dressing as a form of organisational hypocrisy. If the Rating truly reflected stability, the hypotheses would have been confirmed; the opposite pattern raises the possibility that more fragile firms opportunistically obtain the Rating as a masking mechanism.

A key limitation is the absence of a control group of comparable unrated firms. An alternative interpretation is that a longer time horizon may be required for the benefits of the Rating to materialise: a four-year window may be insufficient. In the short term, firms targeting higher scores may incur costs that exceed immediate benefits, which could be interpreted as investments for long-term gains.

5 Conclusion

This study examines the relationship between the Legality Rating and corporate performance using a sample of 1,925 Italian rated firms (7,700 firm-year observations). The results do not support the hypotheses: higher scores are linked to increased OPEX (agency costs) and greater SD(ROA) (operational risk), alongside a decrease in Z_SCORE (default risk). Future research could extend the analysis by including additional dimensions (e.g., earnings management, accruals) and a control sample for comparison. These findings may interest investors and minority shareholders by helping them assess whether corporate actions genuinely enhance reputation or instead serve to mask the expropriation of minority interests.

6 Ethics Declaration

The authors state that no ethical clearance was required for the research presented in this paper. The authors also declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper

7 AI Declaration

During the preparation of this work the authors used Claude AI in order to correct grammar, improve language and ensure readability. After using this tool, the authors reviewed and edited the work and take full responsibility for the content of the publication.

References

Alfano, M.R., Cantabene, C. and de Iudicibus, A. (2024) The effectiveness of a certification of legality: Evidence from Italian firms, MPRA Paper No. 120306. Munich: Munich Personal RePEc Archive.

Altman E. I., Hotchkiss E. (1993). Corporate financial distress and bankruptcy. New York: John Wiley & Sons.

Altman, E.I. (1968) Financial ratios, discriminant analysis and the prediction of corporate bankruptcy, *Journal of Finance*, Vol. 23, No. 4, pp. 589–609.

Ang J. S., Cole R. A., Lin J. W. (2000). Agency costs and ownership structure. The Journal of Finance, vol. 55, n. 1, p. 81-106.

Barnea, A. and Rubin, A. (2010) Corporate social responsibility as a conflict between shareholders, *Journal of Business Ethics*, Vol. 97, No. 1, pp. 71–86.

Berle, A.A. and Means, G.C. (1932) The Modern Corporation and Private Property, Macmillan, New York.

Bertacchini, F., Magri, C. and Gabrielli, G. (2025) What is in a rating? Exploring the link between the Italian Legality Rating and earnings management, *Business Strategy and the Environment*, Vol. 34, No. 1, pp. 1494-1509.

Caputo, A. and Pizzi, S. (2019) Ethical Firms and Web Reporting: Empirical Evidence about the Voluntary Adoption of the Italian "Legality Rating", *International Journal of Business and Management*, Vol. 14, No. 1, pp. 36-45.

Carroll, A.B. (1991) The pyramid of corporate social responsibility: Toward the moral management of organizational stakeholders, *Business Horizons*, Vol. 34, No. 4, pp. 39–48.

Cespa, G. and Cestone, G. (2007) Corporate social responsibility and managerial entrenchment, *Journal of Economics & Management Strategy*, Vol. 16, No. 3, pp. 741–771.

Connelly, B.L., Certo, S.T., Ireland, R.D. and Reutzel, C.R. (2011) Signaling theory: A review and assessment, *Journal of Management*, Vol. 37, No. 1, pp. 39–67.

De Benedetto, M. A., Giuranno, M. G., Mastromarco, C., Ravenda, D., & Toma, P. (2024). Legality rating and corporate efficiency: evidence from a conditional nonparametric frontier analysis. *Empirical Economics*, 66(5), 2137-2168.

Djankov, S., La Porta, R., Lopez-de-Silanes, F. and Shleifer, A. (2008) The law and economics of self-dealing, *Journal of Financial Economics*, Vol. 88, No. 3, pp. 430–465.

Do, T. K. (2022). Corporate social responsibility and default risk: International evidence. *Finance Research Letters*, 44, 102063.

Faccio, M. and Lang, L.H.P. (2002) The ultimate ownership of Western European corporations, *Journal of Financial Economics*, Vol. 65, No. 3, pp. 365–395.

Ferrell, A., Liang, H. and Renneboog, L. (2016) Socially responsible firms, *Journal of Financial Economics*, Vol. 122, No. 3, pp. 585–606.

Freeman, R.E. (1984) Strategic Management: A Stakeholder Approach, Pitman, Boston, MA.

Harjoto M., Laksmana I., (2018). The Impact of Corporate Social Responsibility on Risk Taking and Firm Value. Journal of Business Ethics, Springer, vol. 151, n. 2, p. 353-373.

Jensen, M.C. and Meckling, W.H. (1976) Theory of the firm: Managerial behavior, agency costs and ownership structure, *Journal of Financial Economics*, Vol. 3, No. 4, pp. 305–360.

Jo, H. and Harjoto, M.A. (2011) Corporate governance and firm value: The impact of corporate social responsibility, *Journal of Business Ethics*, Vol. 103, No. 3, pp. 351–383.

La Porta, R., Lopez-de-Silanes, F. and Shleifer, A. (1999) Corporate ownership around the world, *Journal of Finance*, Vol. 54, No. 2, pp. 471–517.

La Porta, R., Lopez-de-Silanes, F., Shleifer, A. and Vishny, R.W. (1997) Legal determinants of external finance, *Journal of Finance*, Vol. 52 No. 3, pp. 1131–1150.

La Porta, R., Lopez-de-Silanes, F., Shleifer, A. and Vishny, R.W. (2000) Investor protection and corporate governance, *Journal of Financial Economics*, Vol. 58, No. 1–2, pp. 3–27.

La Rosa, F. (2020) *Misure di legalità e sviluppo integrale. Analisi delle determinanti e degli effetti del rating legale nelle aziende italiane*, Aracne Editore, Roma.

La Rosa, F., Paternostro, S. and Bernini, F. (2019) What Is Behind the Choice of the Quality of Legality Rating by Italian Private Firms?, in Culasso F., Pizzo, M., Identità, innovazione e impatto dell'aziendalismo italiano. Dentro l'economia digitale, Università di Torino, Torino, pp. 389-400.

La Rosa, F., Paternostro, S. and Bernini, F. (2021) Corporate and regional governance antecedents of the Legality Rating of private Italian companies, *Journal of Management and Governance*, Vol. 27, No.1, pp. 1–33.

Lo Conte, D.L. and Sancetta, G. (2023) Ethical businesses in an ethical world: Does the adoption of the Italian "Legality Rating" affect firms' insolvency risk?, *26th Excellence in Services International Conference - Conference Proceedings*, Glasgow, Scotland, United Kingdom.

Lo Conte, D.L. and Sancetta, G. (2024) Safeguarding Ethics, Sustaining Success: The Role of Italy's Legality Rating in Mitigating Firms' Insolvency Risk,, *Springer Proceedings in Business and Economics, in: Maria Vincenza Ciasullo & Jacques Martin & Federico Brunetti (ed.), Embracing Sustainability Management Through Excellence in Services*, pp. 174-193, Springer.

Maurutto S., (2024). Una, nessuna ... centomila sostenibilità? Dal window dressing alla sostenibilità relazionale e solidale nella gestione della crisi d'impresa, *Ristrutturazione aziendale*, 8 gennaio, pp. 1-27.

McWilliams, A. and Siegel, D. (2001) Corporate social responsibility: A theory of the firm perspective, *Academy of Management Review*, Vol. 26, No. 1, pp. 117–127.

Mulia, S.S. and Joni, J. (2020) Corporate social responsibility and corporate risk-taking: Evidence from an emerging economy, *Journal of Asian Finance, Economics and Business*, Vol. 7, No. 11, pp. 851–858.

Orazalin, N., Mahmood, M. and Narbaev, T. (2019) The impact of sustainability performance indicators on financial stability: Evidence from the Russian oil and gas industry, *Environmental Science and Pollution Research*, Vol. 26, No. 8, pp. 8157–8168.

Orlitzky, M. and Benjamin, J.D. (2001) Corporate social performance and firm risk: A meta-analytic review, *Business & Society*, Vol. 40, No. 4, pp. 369–396.

Pizzi, S., Caputo, A. and Venturelli, A. (2020) Does it pay to be an honest entrepreneur? Addressing the relationship between sustainable development and bankruptcy risk', *Corporate Social Responsibility and Environmental Managemen*t, Vol. 27, No. 3, pp. 1478–1486.

Porter, M.E. and Kramer, M.R. (2002) The competitive advantage of corporate philanthropy, *Harvard Business Review*, Vol. 80, No. 12, pp. 56–68.

Shahrour, M.H., Girerd-Potin, I. and Taramasco, O. (2021) Corporate social responsibility and firm default risk in the Eurozone: A market-based approach, *Managerial Finance*, Vol. 47, No. 7, pp. 975–997.

Spence, M. (1973) Job market signaling, *Quarterly Journal of Economics*, Vol. 87, No. 3, pp. 355–374.

Sun, W. and Cui, K. (2014) Linking corporate social responsibility to firm default risk, *European Management Journal*, Vol. 32, No. 2, pp. 275–287.

Tsoutsoura, M. (2004) Corporate social responsibility and financial performance, Working Paper, Center for Responsible Business, University of California, Berkeley.

Villalonga, B. and Amit, R. (2006) How do family ownership, control and management affect firm value?, *Journal of Financial Economics*, Vol. 80, No. 2, pp. 385–417.

Villalonga, B., Amit, R., Trujillo, M.A. and Guzmán, A. (2015) Governance of family firms, *Annual Review of Financial Economics*, Vol. 7, pp. 635–654.

Weber, M. (2008) The business case for corporate social responsibility: A company-level measurement approach for CSR, *European Management Journal*, Vol. 26, No. 4, pp. 247–261.

Beyond Automation: Strategic Human-Machine Partnerships in Social Enterprise HRM: A Systematic Literature Review

Giuseppe Ambrosio[1], Maria Zifaro[1], Federica Bandini[2], Stefania D'Aprile[1], Giovanni Spatola[1], Marianna Mancino[1]
[1] Universitas Mercatorum, Roma, Italy
[2] Università di Bologna, Bologna, Italy

Abstract

The intersection of artificial intelligence (AI), machine learning technologies, and human resource management (HRM) represents a transformative shift in organisational development. This paper presents a systematic literature review (SLR) examining peer-reviewed evidence on human-machine collaboration in HRM within social enterprises and mission-driven organisations. Using the SPIDER framework and PRISMA guidelines, the review conducted a systematic search on Scopus retrieving 96 peer-reviewed articles, supplemented by backward snowballing.

A critical methodological finding emerged during string construction: incorporating AI-related terms (Cluster C) into the automated TITLE-ABS-KEY search reduced the corpus by 97%, from 245 results to fewer than 12. This reflects a publishing convention whereby AI is discussed substantively in article bodies without being foregrounded in searchable metadata. Cluster C was therefore applied as a manual inclusion criterion during screening rather than as an automated filter.

The review's primary substantive finding is structural: the three-way intersection of AI, HRM, and the third sector is demonstrably scarce in the indexed peer-reviewed literature. The single article meeting all inclusion criteria - Alnamrouti, Rjoub and Ozgit (2022) - provides empirical evidence that AI positively influences sustainable organisational performance in an NGO context, functioning primarily as a mediator amplifying the effect of organisational learning. Backward snowballing on seven references yielded zero additional inclusions, independently corroborating the structural scarcity finding.

Three major research gaps are identified: (1) the absence of domain-specific empirical evidence on AI adoption in social enterprise HRM functions (recruitment, performance management, learning and development, workforce planning, employee engagement); (2) the complete absence of research on ethical and algorithmic risks in mission-driven HRM contexts; and (3) the lack of engagement with established human-machine collaboration frameworks in third-sector management research. These findings establish a documented evidence baseline for a nascent research field, grounding the broader theoretical paper on human-machine partnerships.

Keywords: Human resource management; Social enterprise; Human-machine collaboration; Artificial intelligence; Systematic literature review

Paper type: Academic Research Paper

1 Introduction

The integration of AI into human resource management is reshaping how organisations recruit, develop, and engage their workforces (Malik, Budhwar and Kazmi, 2023; Zhang and Chen, 2024). For social enterprises - organisations defined by the dual imperatives of creating social value while maintaining financial sustainability - this transformation carries distinctive challenges: mission alignment, stakeholder accountability, workforce heterogeneity, and the ethical governance of algorithmic systems operating in contexts where discrimination would directly contradict organisational purpose (Goldkind, Ming and Fink, 2025; Raza Khan et al., 2025).

Despite this theoretical relevance, the peer-reviewed literature has developed in parallel tracks. AI-assisted HRM research concentrates in for-profit settings (Wesche and Sonderegger, 2021; Ruiz, Braojos and Benitez, 2024); social enterprise HRM research is well-developed but largely pre-digital (Oliveira et al., 2021). Their explicit intersection - systematic engagement with how human-machine collaboration is reshaping HRM in mission-driven organisations - is largely absent.

This paper addresses that gap through a systematic literature review. It pursues three research questions:

- *(RQ1) in which HRM domains are social enterprises adopting human-machine collaboration, and what impacts does this generate?*
- *(RQ2) what ethical challenges and algorithmic risks emerge from AI use in mission-driven HRM?*
- *(RQ3) what theoretical frameworks have been mobilised to explain human-machine collaboration in HRM, and what interdisciplinary integration do they evidence?*

The contribution is threefold: the first systematic mapping of peer-reviewed literature at this intersection; empirical documentation of its structural scarcity; and identification of the theoretical and ethical priorities future scholarship must address.

2 Theoretical Background

2.1 Human-machine collaboration beyond automation

Early framings of AI adoption centred on automation - substitution of human labour by systems performing routine tasks more efficiently (Wesche and Sonderegger, 2021). The complementarity perspective offers a more productive account: humans and machines possess systematically different cognitive profiles. Machines excel at processing large data volumes, identifying patterns, and executing consistent procedures without fatigue; humans bring contextual and ethical reasoning, emotional intelligence, and the capacity to navigate value-laden decisions irreducible to optimisation (Jarrahi, 2018; Li et al., 2023). Human-machine collaboration is thus not a transitional state toward full automation but a stable configuration where both jointly accomplish tasks neither could perform as effectively alone (Puranam et al., 2021).

Three frameworks formalise this logic. Wilson and Daugherty (2018) articulate a collaborative intelligence model in which AI augments human capabilities across five dimensions: amplifying cognitive strengths, interacting at scale, embodying skills in physical processes, and enabling new forms of teaming. Jarrahi (2018) develops a human-AI symbiosis perspective emphasising reciprocal learning: AI improves from human feedback while humans develop new competencies. Puranam et al. (2021) situate collaboration within organisational design theory, arguing that task allocation between humans and machines is a deliberate design variable. In HRM, these frameworks converge on a common insight: the value of AI derives not from its capacity to replace human judgement but from its capacity to enhance the quality, consistency, and reach of human decision-making by handling the data-intensive, pattern-recognition, and process-execution dimensions of complex tasks - freeing HR professionals to concentrate on the relational, evaluative, and strategic dimensions of their role that require human sensitivity and contextual understanding (Malik, Budhwar and Kazmi, 2023; Sun and Qiu, 2025).

2.2 AI-Assisted HRM in social enterprises

AI has been applied across five HRM domains relevant to social enterprises. In recruitment and selection, AI screens applications and conducts preliminary interviews via natural language processing (Wesche and Sonderegger, 2021). For resource-constrained social enterprises, this enables processing of larger applicant pools, but human involvement remains essential for assessing mission alignment and ensuring diversity, equity, and inclusion commitments are upheld. In performance management, technology-enabled platforms collect data continuously, provide real-time feedback, and identify competency gaps (Ruiz, Braojos and Benitez, 2024). For social enterprises, systems can integrate multiple dimensions - individual productivity, contribution to social outcomes, and embodiment of organisational values - addressing the subjective biases of traditional appraisal.

In learning and development, adaptive platforms customise content by learning style, pace, and prior knowledge, democratising professional development access (Vardarlier, 2020; Zhang and Chen, 2024). In workforce planning, optimisation algorithms balance operational requirements with employee preferences while predictive models enable proactive resource allocation - critical for social enterprises managing fluctuating demand and volunteer coordination (Malik, Budhwar and Kazmi, 2023). In employee engagement, sentiment analysis tools assess organisational climate continuously. Where commitment to social mission represents a crucial competitive advantage, early detection of emerging tensions enables timely human intervention through authentic relationship-building and values-based dialogue (Ameen et al., 2024). Machine learning can further personalise development pathways for heterogeneous workforces combining paid staff, volunteers, and ex-beneficiaries - while human managers ensure recommendations align with organisational values and individual dignity (Sun and Qiu, 2025).

2.3 Ethical dimensions in mission-driven contexts

Algorithmic bias - systematic unfair differentiation in AI outputs across demographic groups - represents the most consequential risk in HRM adoption (Dima et al., 2024; Raza Khan et al., 2025). For social enterprises, these risks carry distinctive moral weight: organisations whose mission is to promote social inclusion cannot deploy HRM systems that reproduce the inequalities they exist to address without entering into fundamental contradiction with their institutional purpose (Goldkind, Ming and Fink, 2025). Effective governance requires

algorithmic auditing, human override mechanisms, transparent communication about AI use, and proportionate accountability structures - designed to be adequate to the ethical stakes while respecting resource constraints.

3 Methods

3.1 Review design

This study adopts a systematic literature review (SLR) methodology following PRISMA guidelines. Research questions were operationalised using the SPIDER framework (Cooke, Smith and Booth, 2012) which captures qualitative, quantitative, and theoretical contributions across heterogeneous study designs, appropriate for complex organisational phenomena. Table 1 presents the full operationalisation.

Table 1: SPIDER framework operationalisation

SPIDER	RQ1 - Adoption & impacts	RQ2 - Ethical dimensions	RQ3 - Theoretical frameworks
S - Sample	Social enterprises, nonprofits, third sector	Mission-driven orgs, HR stakeholders	Interdisciplinary academic literature
PI - Phenomenon	AI/ML in HRM + performance effects	Algorithmic bias, fairness, AI governance	Theories of human-machine collaboration
D - Design	Case study, survey, mixed-methods	Critical analysis, normative frameworks	Literature review, bibliometrics
E - Evaluation	Efficiency, engagement, social outcomes	Discriminatory risks, mitigation strategies	Dominant theories, theoretical gaps
R - Research type	Empirical	Theoretical, normative, empirical	Theoretical, bibliometric

Source: Authors' elaboration

3.2 Search strategy

The bibliographic search was conducted on Scopus. The string combined Cluster A (organisational context: social enterprises, nonprofits) and Cluster B (HRM domain: human resource management, talent management). AI and human-machine collaboration terms (Cluster C) were deliberately excluded from the automated string and applied as a manual inclusion criterion during screening.

This choice was motivated by systematic diagnostic testing: 14 iterative tests showed that incorporating Cluster C into TITLE-ABS-KEY reduces the corpus by approximately 97% (from 245 to fewer than 12 results). This reflects a publishing convention - AI is treated substantively in article bodies without being foregrounded in searchable metadata. Embedding Cluster C automatically would systematically exclude pertinent contributions. The final string was:

TITLE-ABS-KEY (("social enterprise" OR "nonprofit*" OR "non-profit organization*") AND ("human resource management" OR "HRM" OR "talent management")) AND DOCTYPE (ar) AND LANGUAGE (english) AND PUBYEAR > 2020*

String parameters (peer-reviewed articles, English, published 2021+) automatically satisfied inclusion criteria for document type, language, and publication period.

3.3 Inclusion and exclusion criteria

Three manual inclusion criteria were applied sequentially during screening (I1 → I2 → I3). Criteria for document type, publication period, and language were pre-empted by the search string. Table 2 presents both inclusion and exclusion criteria.

Table 2: Manual screening criteria (I1–I3 and E1–E3)

Code	Criterion	Description
I1	Organisational context	Social enterprises, nonprofits, not-for-profit, third sector or mission-driven organisations
I2	HRM domain	At least one HRM area: recruitment, performance management, talent management, workforce planning, employee engagement, or L&D
I3 *	AI / human-machine collab.	Substantive engagement with AI, ML, algorithms, automation, or human-machine collaboration in the HRM context

Code	Criterion	Description
E1	Non-pertinent context	Profit-oriented, public administration, or industrial/manufacturing organisations
E2	HRM absent	No HR management dimensions - focus on finance, marketing, or operations
E3	AI absent or marginal	AI or digital technology mentioned only marginally without substantial analytical contribution

** I3 = Cluster C applied manually. Criteria for document type (peer-reviewed), period (2021+) and language (English) are guaranteed by string parameters.*
Source: Authors' elaboration

3.4 Screening and Snowballing

The 96 retrieved records were imported into Rayyan. Duplicate detection confirmed zero duplicates. Title, abstract, and keyword screening applied I1–I3 sequentially. Two articles satisfied I1 and I2; full-text assessment confirmed one - Alnamrouti, Rjoub and Ozgit (2022) - met all three criteria. Lin, Lu and Sun (2023) was excluded under E3: its use of data analytics in a blood bank staffing context did not constitute AI or human-machine collaboration substantively. The 31 'Maybe' records - pertinent to context and HRM but lacking technological signal in metadata - were excluded under E3.

Backward snowballing following Wohlin (2014) analysed seven references from Alnamrouti et al. (2022). None met all inclusion criteria (Table 3). Forward snowballing was not conducted: backward snowballing corroborated the structural scarcity independently, and the expected yield of articles citing a 2022 paper at the exact three-way intersection was assessed as negligible.

Table 3: Backward snowballing results - Alnamrouti et al. (2022)

#	Reference	Verdict	Exclusion criterion
1	Wamba-Taguimdje et al. (2020) - AI influence on firm performance. Business Process Management Journal	Excluded	E1 + E2
2	Hamadamin & Atan (2019) - SHRM and competitive advantage. Sustainability	Excluded	E1 + E3
3	Al Aina & Atan (2020) - Talent Management and SOP. Sustainability	Excluded	E1 + E3
4	Mikalef & Gupta (2021) - AI capability. Information & Management	Excluded	E1
5	Tariq et al. (2021) - Operational Excellence Through AI. Frontiers in Psychology	Excluded	E1
6	Sarker & Rahman (2018) - SHRM Practices in NGOs. European Journal of Business and Management	Excluded	E3
7	Oliveira et al. (2021) - Strategy and HRM in Non-Profits. Journal of Open Innovation	Excluded	E3

Source: Authors' elaboration

Data extraction from the included article used a structured grid organised around the three RQs. Given the corpus size of one article, a narrative synthesis approach was adopted consistent with best practice for limited corpora (Tranfield, Denyer and Smart, 2003).

4 Results and Discussion

4.1 Search results

The Scopus search returned 96 articles; no duplicates were found. Screening identified two articles satisfying I1 and I2; full-text assessment confirmed one as meeting all three inclusion criteria. Backward snowballing on seven references yielded zero additional inclusions. Figure 1 [PRISMA flow diagram - see attached] illustrates the complete screening process.

4.2 Structural scarcity as a primary finding

The most significant finding emerges from the search process itself: the three-way intersection of AI/human-machine collaboration, HRM, and the third sector is structurally scarce. The HRM-in-third-sector literature (A+B: 245 results) and the HRM-and-AI literature (B+C: 3,406 results) are individually substantial; their three-way intersection collapses by 97%. Backward snowballing corroborates this independently: zero of seven references from the sole included article satisfies the intersection. This scarcity is a feature of the literature, not a methodological artefact, and constitutes original empirical evidence in its own right - signalling a consequential blind spot in management and organisation research.

4.3 Characteristics of the included study

Alnamrouti, Rjoub and Ozgit (2022) conducted an empirical quantitative study in Jordan, drawing on 171 HR professionals from 21 United Nations agencies and 70 NGOs. Using PLS-SEM (WarpPLS 7.0), they tested a structural model in which organisational learning (OL) and corporate social responsibility (CSR) predict sustainable organisational performance (SOP), mediated by strategic HRM (SHRM) and AI. Table 4 presents the structured data extraction.

Table 4: Data extraction - Alnamrouti, Rjoub and Ozgit (2022)

Dimension	Alnamrouti, Rjoub & Ozgit (2022)
Journal	Sustainability (MDPI), 14, 7327
Country / context	Jordan - 21 UN agencies + 70 NGOs
Method	Quantitative survey - PLS-SEM (WarpPLS 7.0), n = 171
AI construct	6-item scale - AI as general organisational capability
Key finding RQ1	AI → SOP: β = 0.13, p = 0.04 (direct effect); OL → AI → SOP: β = 0.56, p = 0.03 (strongest mediation path)
Key finding RQ2	Not addressed. Sole normative note: recommendation to establish a 'control tower for data ethics'
Key finding RQ3	RBV + OL Theory + IT Capability Theory. No engagement with human-machine collaboration frameworks
Model fit	GOF = 0.667 (large); SRMR = 0.096 (acceptable); R^2 SOP = 0.09 (low)
Main limitation	Non-probability sampling; single sector (Jordanian NGOs); AI as aggregate construct - no HRM domain specificity

Source: Authors' elaboration from full-text analysis

4.4 Findings by research question

RQ1 - Adoption and Impacts

Alnamrouti et al. (2022) provide the only peer-reviewed empirical evidence that AI positively influences organisational performance in a nonprofit context. AI functions as a direct predictor of SOP (β = 0.13, p = 0.04) and, more substantially, as a mediator amplifying organisational learning's effect on performance (β = 0.56, p = 0.03 - the model's strongest path). However, AI is treated as an aggregate capability construct: no empirical evidence addresses specific HRM domain applications in social enterprises. The mediation result is consistent with the complementarity perspective (Wilson and Daugherty, 2018; Puranam et al., 2021), though this connection is not developed by the authors. The low explained variance (R^2 = 0.09) suggests additional variables - including domain-specific human-machine collaboration constructs - may be relevant.

RQ2 - Ethical Dimensions

The included study does not address algorithmic bias, fairness, data governance, or ethical implications of AI in HRM. The sole normative indication is a brief recommendation to establish a 'control tower for data ethics'. This silence is representative of a broader structural gap: the AI ethics in HRM literature (Dima et al., 2024; Raza Khan et al., 2025) has developed independently of the nonprofit and social enterprise context. The ethical risks well-documented in for-profit settings - discriminatory screening, biased performance appraisal, opacity in automated decisions - remain empirically uninvestigated in organisations whose mission is to counter precisely the inequalities biased algorithms can reproduce.

RQ3 - Theoretical Frameworks

The included study mobilises Resource-Based View (Barney, 1991), Organisational Learning Theory (Argyris and Schon, 1978), and IT Capability Theory (Wamba-Taguimdje et al., 2020). This grounding is appropriate for performance-oriented quantitative modelling but does not engage with frameworks developed specifically for human-machine collaboration - the complementarity perspective (Puranam et al., 2021), the human-AI symbiosis model (Jarrahi, 2018), or collaborative intelligence (Wilson and Daugherty, 2018). The absence of these frameworks from the third-sector HRM literature represents a theoretical gap and the clearest opportunity for the present paper's contribution.

5 Conclusions

5.1 Summary and theoretical contributions

This systematic review mapped peer-reviewed evidence on human-machine collaboration in HRM within social enterprises. The primary result is structural: one study meets all inclusion criteria. Three theoretical

contributions follow. First, this review provides systematic evidence that the intersection of AI, HRM, and the third sector constitutes a largely uncharted research territory. Prior work has documented the rapid development of AI-assisted HRM in for-profit organisations (Malik, Budhwar and Kazmi, 2023; Zhang and Chen, 2024) and has separately examined HRM challenges specific to social enterprises and nonprofits (Barenblat and Gosselink, 2024; Goldkind, Ming and Fink, 2025), but the two streams have not converged. The present review documents this non-convergence systematically, establishing a baseline against which future research accumulation can be measured. Second, the review reveals a theoretical mismatch between the frameworks available in the literature and the phenomenon under investigation. The frameworks mobilised in the sole included study - RBV, OL theory, IT capability theory - are well-suited to performance-oriented quantitative modelling but do not capture the complementarity dynamics, ethical tensions, and mission-alignment challenges that characterise human-machine partnerships in organisations driven by social purpose. The collaborative intelligence model (Wilson and Daugherty, 2018), the human-AI symbiosis perspective (Jarrahi, 2018), and the human-machine teaming literature (Puranam et al., 2021) offer conceptual resources that have yet to be applied in this context. Their integration with theories of social entrepreneurship and stakeholder management represents a productive direction for future theoretical development. Third, the review establishes that the ethical dimension of AI in social enterprise HRM is not merely understudied but absent from the empirical literature. This absence is particularly significant because the risks associated with algorithmic decision-making in HRM - discriminatory screening, biased performance appraisal, opacity in automated decisions - are well-documented in for-profit contexts (Dima et al., 2024; Raza Khan et al., 2025) but have never been examined in organisations whose explicit mission is to counter precisely the forms of inequality that biased algorithms can reproduce. Establishing an empirical and normative agenda for AI ethics in social enterprise HRM is therefore both theoretically urgent and practically consequential.

5.2 Practical Implications and Limitations

For organisational leaders, the review signals that evidence-based guidance on AI adoption in nonprofit HRM does not yet exist at domain-specific level. Tools validated in for-profit settings carry assumptions about performance metrics and workforce composition that may not translate. The establishment of internal governance mechanisms - including algorithmic auditing and transparent employee communication - is advisable regardless of external evidence maturity. For policymakers, structural scarcity points to urgent need for sector-specific research investment: the rapid diffusion of AI-powered HRM tools (Barenblat and Gosselink, 2024; Brine et al., 2025; Valenti, 2025) is proceeding without peer-reviewed evidence on effectiveness, equity, or mission alignment.

This review is limited to a single database (Scopus) and to English-language articles published from 2021 onwards. The restriction to English is a potentially significant limitation given the geographical concentration of social enterprise research in non-Anglophone European countries, including Italy, France, and Spain - precisely the contexts most relevant to the phenomenon under investigation. The corpus of one article precludes meta-analytic synthesis - a constraint reflecting the state of the field, not the methodology. Future priorities include domain-specific empirical studies on AI in social enterprise HRM (particularly in the Italian and European context given the sector's institutional maturity), empirical investigation of algorithmic risk in nonprofit HRM, and theoretical integration of human-machine collaboration frameworks with social enterprise theory and institutional logics.

6 Ethics Declaration

This research did not required official ethical approvement

7 AI Declaration

This paper was developed with the support of Claude (Anthropic), an AI assistant, which was specifically used to: (1) support the diagnostic testing of the Scopus search string; (2) support the data extraction; and (3) assist in revising the paper in english. All methodological decisions - including the choice of search strategy, inclusion and exclusion criteria, the decision not to conduct forward snowballing, and all analytical interpretations - were made exclusively by the authors.

References

Al Aina, R. and Atan, T. (2020) 'The impact of implementing talent management practices on sustainable organizational performance', Sustainability, Vol. 12, p. 8372.

Alnamrouti, A., Rjoub, H. and Ozgit, H. (2022) 'Do strategic human resources and artificial intelligence help to make organisations more sustainable? Evidence from non-governmental organisations', Sustainability, Vol. 14, p. 7327.
Ameen, N. et al. (2024) 'The rise of human-machine collaboration: Managers' perceptions of leveraging AI for enhanced B2B service recovery', British Journal of Management, Vol. 35, No. 2, pp. 245-268.
Argyris, C. and Schon, D. (1978) Organizational Learning: A Theory of Action Perspective, Addison-Wesley, Reading MA.
Barenblat, K. and Gosselink, B.H. (2024) 'Mapping the landscape of AI-powered nonprofits', Stanford Social Innovation Review, Summer 2024.
Barney, J. (1991) 'Firm resources and sustained competitive advantage', Journal of Management, Vol. 17, No. 1, pp. 99-120.
Brine, D., Wong, N., Porway, J. and Porteous, R. (2025) AI Can't Be Ignored: Exploring the Opportunities for Nonprofits and the Social Sector, Bridgespan Group [online] https://www.bridgespan.org/insights/exploring-ai-opportunities-for-nonprofits-and-the-social-sector
Cooke, A., Smith, D. and Booth, A. (2012) 'Beyond PICO: The SPIDER tool for qualitative evidence synthesis', Qualitative Health Research, Vol. 22, No. 10, pp. 1435-1443.
Deloitte Insights (2025) Strengthening the Bonds of Human and Machine Collaboration [online] https://www.deloitte.com/us/en/insights/topics/talent/human-machine-collaboration.html
Dima, J., Gilbert, M.-H., Dextras-Gauthier, J. and Giraud, L. (2024) 'The effects of AI on human resource activities and the roles of the human resource triad', Frontiers in Psychology, Vol. 15.
Goldkind, L., Ming, J. and Fink, A. (2025) 'AI in the nonprofit human services: Distinguishing between hype, harm, and hope', Human Service Organizations: Management, Leadership & Governance, Vol. 49.
Hamadamin, H.H. and Atan, T. (2019) 'The impact of strategic human resource management practices on competitive advantage sustainability', Sustainability, Vol. 11, p. 5782.
Jarrahi, M.H. (2018) 'Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making', Business Horizons, Vol. 61, No. 4, pp. 577-586.
Li, J.-M., Wu, T.-J., Wu, Y.J. and Goh, M. (2023) 'Systematic literature review of human-machine collaboration in organizations using bibliometric analysis', Management Decision, Vol. 61, No. 1, pp. 45-75.
Lin, J., Lu, M. and Sun, P. (2023) 'Worker experience and donor heterogeneity: The impact of charitable workers on donors' blood donation decisions', Manufacturing & Service Operations Management.
Malik, A., Budhwar, P. and Kazmi, B.A. (2023) 'AI-assisted HRM: Towards an extended strategic framework', Human Resource Management Review, Vol. 33, No. 1, p. 100940.
Masum, A.K.M., Beh, L.-S., Azad, M.A.K. and Hoque, K. (2018) 'Intelligent human resource information system (i-HRIS): A holistic decision support framework for HR excellence', International Arab Journal of Information Technology, Vol. 15, No. 1, pp. 121-130.
Mikalef, P. and Gupta, M. (2021) 'Artificial intelligence capability: Conceptualisation, measurement calibration, and empirical study on its impact on organisational creativity and firm performance', Information & Management, Vol. 58, No. 3, p. 103434.
Oliveira, M., Sousa, M., Silva, R. and Santos, T. (2021) 'Strategy and human resources management in non-profit organisations: Its interaction with open innovation', Journal of Open Innovation: Technology, Market, and Complexity, Vol. 7, p. 75.
Puranam, P., Shrestha, Y.R., He, V.F. and von Krogh, G. (2021) 'Human-machine collaboration in managerial decision making', Computers in Human Behavior, Vol. 119, p. 106722.
Raza Khan, S., Baig, M., Qaiser, A., Abrar, M. and Ali, H. (2025) 'Humans and machines at work: Redefining strategic HRM in the era of AI and workforce automation', The Critical Review of Social Sciences Studies, Vol. 3, No. 2, pp. 145-172.
Ruiz, L., Braojos, J. and Benitez, J. (2024) 'Digital human resource strategy: Conceptualization, theoretical development, and empirical examination', Information & Management, Vol. 61, No. 4.
Sarker, M.S.I. and Rahman, M.A. (2018) 'Effect of strategic HRM practices on perceived financial performance of non-governmental organisations', European Journal of Business and Management, Vol. 10, pp. 99-110.
Sun, J. and Qiu, L. (2025) 'Human-machine collaboration: Definitions, models, and socio-economic impacts', International Journal of Crowd Science, Vol. 9, No. 1, pp. 1-24.
Tariq, M.U., Poulin, M. and Abonamah, A.A. (2021) 'Achieving operational excellence through artificial intelligence: Driving forces and barriers', Frontiers in Psychology, Vol. 12, p. 686624.
Tranfield, D., Denyer, D. and Smart, P. (2003) 'Towards a methodology for developing evidence-informed management knowledge by means of systematic review', British Journal of Management, Vol. 14, No. 3, pp. 207-222.
Valenti, K. (2025) How Nonprofits Are Leveraging AI for Social Impact in 2025, Sigma Forces [online] https://www.sigmaforces.com/post/nonprofits-ai-social-impact-2025
Vardarlier, P. (2020) 'Digital transformation of human resource management', in Hacioglu, U. (Ed.) Digital Business Strategies in Blockchain Ecosystems, Springer, pp. 239-264.
Wamba-Taguimdje, S.L., Wamba, S.F., Kamdjoug, J.R.K. and Wanko, C.E.T. (2020) 'Influence of artificial intelligence (AI) on firm performance: The business value of AI-based transformation projects', Business Process Management Journal, Vol. 26, No. 7, pp. 1893-1924.
Wesche, J.S. and Sonderegger, A. (2021) 'When computers take the lead: The automation of leadership', Computers in Human Behavior, Vol. 101, pp. 197-209.
Wilson, H.J. and Daugherty, P.R. (2018) 'Collaborative intelligence: Humans and AI are joining forces', Harvard Business Review, Vol. 96, No. 4, pp. 114-123.

Wirtz, B.W., Weyerer, J.C. and Geyer, C. (2019) 'Artificial intelligence and the public sector: Applications and challenges', International Journal of Public Administration, Vol. 42, No. 7, pp. 596-615.

Wohlin, C. (2014) 'Guidelines for snowballing in systematic literature studies and a replication in software engineering', Proceedings of EASE 2014, ACM.

Zhang, J. and Chen, Z. (2024) 'Exploring human resource management digital transformation in the digital age', Journal of the Knowledge Economy, Vol. 15, No. 1, pp. 1482-1498.

The Italian Crisis and Insolvency Code (CCII) Adequate Organisational Frameworks and AI Predictive Judgments

Domenico Celenza[1], Marcella Caradonna[2], Francesca Francioli[3], Stefano Gazzani[1]

[1] University of Cassino and Southern Lazio, Cassino, Italy
[2] Catholic University of Milan, Milano, Italy
[3] University of Genoa, Genoa, Italy

Abstract

The Italian Crisis and Insolvency Code has moved corporate governance from a reactive idea of crisis management to a preventive model in which firms are expected to organise information, responsibilities and controls before distress becomes irreversible. This paper develops a conceptual framework for interpreting adequate organisational, administrative and accounting arrangements as an intelligent and sustainable governance infrastructure. It argues that Artificial Intelligence (AI) can enhance such arrangements not simply by improving insolvency prediction, but by transforming dispersed organisational data into shared knowledge for timely, accountable and human-centred decisions. The analysis is positioned at the intersection of the knowledge-based theory of the firm, management control, corporate governance and the literature on algorithmic decision support. It considers adequate arrangements as dynamic knowledge architectures that connect monitoring, interpretation and intervention, and it examines how predictive analytics, machine learning and explainable AI may support early warning systems, business continuity and organisational resilience. The contribution of the paper is threefold. First, it reframes legal adequacy as a managerial and knowledge-management capability, rather than as a formal compliance exercise. Second, it proposes an AI-enabled model in which data governance, predictive judgement, human oversight and ethical accountability are treated as mutually dependent elements of the same control environment. Third, it links crisis prevention to the broader agenda of intelligent and sustainable organisations, showing that predictive systems can help boards and managers reconcile continuity, stakeholder protection and sustainable value creation. The paper also discusses critical risks associated with AI adoption, including model opacity, poor data quality, bias, cybersecurity vulnerability and excessive delegation of judgement to automated systems. These risks require boards to preserve interpretive responsibility and to build validation, auditability and learning routines into the design of intelligent arrangements. The paper concludes that AI may strengthen the practical effectiveness of adequate arrangements only when embedded in relational governance, organisational learning and explicit ethical safeguards. Its main implication is that the future of crisis governance depends less on technological sophistication alone than on the capacity of organisations to combine human and artificial intelligence in a coherent knowledge system.

Keywords: corporate governance, adequate organisational arrangements, insolvency prediction, artificial intelligence, business continuity.

Paper type: Academic Research Paper

1 Introduction

The future of sustainable organisations depends on their capacity to transform intelligence into responsible action. In turbulent environments, intelligence cannot be reduced to the possession of data or to the adoption of advanced digital tools. It concerns the ability to notice weak signals, understand their meaning, share them across organisational boundaries and translate them into timely decisions. This problem is particularly visible in crisis governance. Financial distress is rarely a sudden event. It is normally preceded by declining margins, tensions in working capital, delays in payments, deteriorating relations with banks or suppliers, and managerial hesitation in acknowledging unfavourable signals. The challenge is therefore not only to predict crisis, but to organise knowledge before crisis disables the organisation's capacity to act.

The Italian Crisis and Insolvency Code, introduced by Legislative Decree No. 14/2019 and subsequently amended, gives legal force to this preventive orientation. It requires entrepreneurs and corporate bodies to adopt measures and arrangements capable of detecting the state of crisis promptly and of supporting action for business continuity. In this respect, the Code should be read not only as an insolvency reform, but also as a governance reform. It shifts attention from ex post allocation of losses to ex ante organisation of information flows, responsibilities and controls. The duty to establish adequate organisational, administrative and accounting arrangements asks the firm to become intelligible to itself. It requires an internal architecture through which dispersed operational events can be converted into reliable managerial knowledge.

This paper develops the argument that adequate arrangements are a specific form of intelligent organisation. They are intelligent because they connect sensing, interpretation and response; they are sustainable because their purpose is not mere short-term survival, but the preservation of business continuity, stakeholder confidence and value creation. The topic is aligned with the debate on intelligent and sustainable communities and organisations, where human, collective and artificial intelligence interact to support innovation, resilience and inclusive growth. The firm that prevents crisis through better knowledge does not only protect shareholders. It protects employees, creditors, local communities, supply chains and public confidence in economic activity.

Artificial Intelligence (AI) offers a powerful but ambivalent contribution to this transition. Machine learning techniques can process large quantities of accounting, operational and contextual data, detect non-linear patterns and generate early warnings that traditional ratio-based models may miss. AI can reduce information asymmetries and support boards in the continuous monitoring of business continuity. However, algorithmic prediction is not equivalent to managerial judgement. A model may classify a firm as risky without explaining why; it may reproduce biases contained in historical data; it may appear objective while hiding fragile assumptions; it may encourage excessive reliance on automated outputs. Consequently, AI can strengthen adequate arrangements only if it is embedded in a governance system that preserves human accountability, transparency, validation and organisational learning.

The research question guiding the paper is therefore: how can AI be integrated into adequate organisational, administrative and accounting arrangements so as to support sustainable crisis governance and knowledge-driven value creation? The paper adopts a conceptual research design. It synthesises legal and managerial interpretations of adequate arrangements with knowledge-management theory, insolvency prediction literature and AI governance studies. The aim is not to test a predictive model, but to propose a coherent framework for understanding the role of AI within the governance of continuity.

The remainder of the paper is organised as follows. The next section discusses adequate arrangements through the lens of management control and knowledge-based theory. The third section examines AI-enabled insolvency prediction and its limits. The fourth section proposes a conceptual architecture for AI-enabled adequate arrangements. The fifth section discusses implications for boards, control bodies and practitioners. The final section summarises contributions, limitations and future research directions.

2 Adequate Arrangements as Knowledge Infrastructure

In the Italian aziendalistica tradition, organisational adequacy is not a purely formal requirement. It concerns the correspondence between organisational structures, administrative procedures, accounting systems and the complexity of the firm. An arrangement is adequate when it allows the organisation to function coherently, when it clarifies responsibilities, when it supports planning and control, and when it produces information that is sufficiently reliable, timely and relevant for decision-making. From this perspective, adequate arrangements are close to the classical idea of economic order within the firm: they make action possible because they stabilise roles, connect activities and create a common informational basis.

The Crisis and Insolvency Code reinforces this interpretation by linking adequacy to the prompt detection of crisis and to the preservation of business continuity. The legal obligation is therefore inseparable from managerial capability. It is not enough to possess accounting records or periodic reports. The firm must be able to recognise emerging imbalances, evaluate their causes, identify available responses and document the reasoning that leads to action or inaction. Adequacy is a functional criterion: the same tool may be sufficient in a small, simple business and inadequate in a complex group exposed to volatile markets, international supply chains and financial leverage.

This functional reading connects the legal framework with the knowledge-based theory of the firm. Grant (1996) argues that firms exist because they integrate specialised knowledge that individuals hold separately. Nonaka and Takeuchi (1995) describe knowledge creation as a dynamic process in which tacit and explicit knowledge interact. Applied to crisis prevention, these theories show that distress signals do not become useful simply because they exist in databases. They become useful when they are interpreted, shared, challenged and incorporated into decisions. A late payment from a strategic customer may be an isolated event, a symptom of sectoral stress or the beginning of a liquidity shortage. The meaning depends on the organisation's capacity to combine accounting data, commercial information, operational experience and external context.

Adequate arrangements can therefore be understood as knowledge infrastructures. They define what information is collected, who is responsible for it, how often it is updated, how it is validated, who receives it and how it enters the decision process. They also shape the relationship between formal data and practical judgement. In many firms, relevant knowledge is local: sales staff perceive deteriorating customer behaviour, purchasing managers see supplier tensions, financial officers notice covenant pressure, and operational

managers observe productivity losses. If these signals remain fragmented, governance bodies receive information too late or in a form that does not support action. The purpose of adequate arrangements is to avoid this fragmentation.

The sustainability dimension arises because crisis prevention protects the conditions for long-term value creation. Business continuity is not only a financial concept. It involves the capacity to maintain employment, fulfil obligations, preserve knowledge assets, protect relationships with stakeholders and avoid the destruction of organisational capabilities. Sustainable governance therefore requires a broader view of performance in which liquidity, profitability, resilience, environmental and social impacts are observed together. The firm must not become sustainable after crisis; it must organise itself sustainably in order to avoid crisis or to manage it without unnecessary loss of value.

3 Artificial Intelligence and Predictive Crisis Governance

Insolvency prediction has a long history in accounting and finance. Beaver (1966) showed that financial ratios could anticipate failure, and Altman (1968) developed a discriminant model that became a reference point for bankruptcy prediction. Later models, including Ohlson's logit approach and Shumway's hazard model, improved statistical treatment and incorporated the dynamic nature of failure risk. These contributions remain important because they clarify the informational content of accounting and market variables. Yet traditional models also display limits. They often assume relatively stable relationships among variables, rely on restricted financial indicators and struggle to capture qualitative signals, structural breaks and non-linear interactions.

AI-based models respond to some of these limits. Machine learning algorithms can analyse larger and more heterogeneous datasets, including accounting ratios, cash-flow indicators, payment behaviour, tax and social security exposures, operational performance, sectoral trends, supply-chain disruptions and macroeconomic conditions. They can identify interactions that are difficult to specify ex ante and can update predictions as new information becomes available. In a governance context, this may improve the timeliness of early warnings and reduce false negatives, which are particularly dangerous because they lead boards to underestimate emerging distress.

However, predictive accuracy is only one dimension of crisis governance. A model that predicts well but cannot be explained may be difficult to use in board deliberations, internal audit reviews or negotiations with stakeholders. Directors need to know which variables drive the alert, whether the result is robust, how it compares with managerial evidence and what actions are available. Inadequate explanations may turn AI into a black box that increases, rather than reduces, uncertainty. Explainable AI techniques, such as local explanation methods and interpretable models, can help by showing the contribution of specific variables to a prediction. Still, explanation is not merely a technical add-on; it is part of the governance process through which a prediction becomes accountable knowledge.

AI also changes the temporal structure of control. Traditional control systems are often periodic. They compare actual results with budgets, analyse deviations and prepare reports at monthly or quarterly intervals. Predictive systems can operate more continuously. They can monitor cash conversion cycles, customer concentration, cost inflation, covenant headroom and risk indicators in near real time. This continuous orientation is useful only if the organisation defines thresholds, escalation rules and responsibilities. Otherwise, the multiplication of alerts may produce noise, alert fatigue or defensive behaviour. The intelligence of the system depends on the relationship between algorithmic detection and organisational response.

The adoption of AI also introduces new risks. Data quality is central: incomplete, inconsistent or biased data generate misleading outputs. Model risk is equally important: algorithms trained on past failures may not perform well in new economic conditions or in sectors with specific characteristics. Cybersecurity and access controls matter because predictive systems often concentrate sensitive information. Finally, overreliance on AI may weaken managerial responsibility. If directors treat the model as a substitute for judgement, the organisation may become technologically advanced but governantly weak. The appropriate objective is therefore not automation of crisis governance, but augmentation of human and collective intelligence.

4 Research Design and Conceptual Synthesis

This paper is conceptual and theory-building in nature. It does not present an empirical test or a new algorithm. Instead, it integrates four streams of literature and practice: the Italian debate on adequate organisational, administrative and accounting arrangements; knowledge-management theory; insolvency prediction research; and AI governance. The synthesis is justified by the fact that each stream captures only part of the problem. Legal and accounting literature clarifies duties, controls and continuity. Knowledge-management theory explains how information becomes organisational knowledge. Insolvency prediction

studies focus on the anticipation of failure. AI governance addresses transparency, accountability and the responsible use of algorithmic systems.

The conceptual logic is based on three assumptions. First, crisis prevention is a knowledge process. It begins with data collection but requires interpretation, communication and decision. Second, AI is an enabling technology rather than an autonomous governance subject. Its value depends on organisational routines, data governance and human oversight. Third, sustainability and resilience are not external additions to crisis governance. They are part of the same managerial problem because a firm that fails to preserve continuity destroys economic, social and relational value.

On this basis, the paper proposes a framework in which adequate arrangements are organised around four connected layers: data and knowledge governance, predictive analytics, human oversight, and sustainable response. The first layer concerns the quality, ownership and integration of information. The second layer concerns the analytical tools used to detect risks and develop scenarios. The third layer concerns the deliberative mechanisms through which boards, executives and control bodies interpret AI outputs. The fourth layer concerns the actions taken to preserve continuity and stakeholder value. The four layers are mutually dependent. A sophisticated algorithm is ineffective without reliable data; reliable data are insufficient without interpretation; interpretation has limited value without timely intervention; intervention is fragile if it ignores stakeholder and sustainability effects.

This architecture also allows a more precise understanding of the term predictive judgement. Prediction is the statistical estimation of possible outcomes. Judgement is the managerial and governance act of interpreting those outcomes in context. AI can support prediction, but judgement remains a human and collective responsibility. Adequate arrangements must therefore be designed to connect prediction and judgement without confusing them. The board should not ask whether the model has decided that the firm is in crisis. It should ask what the model indicates, why it indicates it, what alternative evidence exists, which assumptions are critical and what actions are proportionate.

5 A Framework for AI-Enabled Adequate Arrangements

An AI-enabled adequate arrangement can be described as an integrated system that transforms organisational signals into accountable decisions for continuity and sustainability. The system begins with data governance. Relevant data are not limited to annual financial statements. They include management accounts, cash forecasts, overdue receivables, supplier payment terms, inventory rotation, customer concentration, order backlog, bank covenants, tax liabilities, litigation risks, employee turnover and external indicators. The organisation must define who owns each data source, how data are verified, how inconsistencies are resolved and how updates are documented. Without this foundation, AI amplifies disorder rather than intelligence.

The second component is predictive analytics. Algorithms may classify risk, estimate probability of default, simulate liquidity scenarios or detect anomalies in accounting and operational flows. Different tools can coexist. Transparent scorecards may support communication with boards and smaller firms, while more complex models may be used for scenario analysis in larger organisations. The key is proportionality. Adequacy does not require every firm to adopt the most sophisticated AI system. It requires a system appropriate to size, sector, risk profile and organisational complexity. In many cases, a hybrid architecture combining rule-based indicators, statistical models and managerial dashboards may be more useful than an opaque algorithm claiming high accuracy.

The third component is interpretive governance. AI outputs should enter a structured decision process. Alerts must be assigned to responsible managers; material risks must be escalated to directors and control bodies; the reasons for accepting, revising or rejecting an alert must be documented. Internal audit or equivalent control functions should periodically test data quality, model performance and compliance with governance procedures. The board should receive information in a form that is concise but not simplistic: risk scores, drivers, sensitivity analysis, trend changes and suggested questions for deliberation. Such practices transform AI from a technical application into a knowledge-management routine.

The fourth component is ethical and legal accountability. The European regulatory environment increasingly emphasises transparency, risk management and human oversight in AI systems. Even where a crisis-prediction tool does not fall into the most restrictive legal categories, firms should apply prudent governance principles because the consequences of erroneous predictions can be serious. A false positive may damage reputation, influence credit relations or trigger unnecessary restructuring. A false negative may delay intervention and increase losses. Responsible AI use therefore requires clear documentation of model purpose, training data, validation results, limitations, access rights and escalation procedures.

The fifth component is sustainable response. The purpose of early warning is not to produce anxiety but to create time for action. If an AI system identifies liquidity deterioration, the organisation should be able to revise

cash planning, negotiate with financial partners, adjust investment timing, review cost structures, protect critical knowledge assets and communicate with stakeholders when appropriate. The response should not be limited to financial repair. It should consider operational resilience, employee implications, supplier continuity, environmental commitments and the reputational effects of managerial choices. In this way, AI-enabled arrangements support both crisis prevention and sustainable value creation.

Table 1: A conceptual architecture for AI-enabled adequate arrangements

Layer	Knowledge function	AI contribution	Governance safeguard
Data governance	Creates a reliable common base for interpretation	Integrates accounting, operational and contextual data	Ownership, validation, access rules and audit trails
Predictive analytics	Transforms signals into risk scenarios	Detects patterns, anomalies and probability of distress	Model validation, proportionality and explainability
Human oversight	Connects prediction with deliberation	Provides dashboards and decision support	Board accountability, escalation rules and documented judgement
Sustainable response	Converts warnings into continuity actions	Supports scenario planning and resource prioritisation	Stakeholder assessment and review of outcomes

6 Discussion

The proposed framework has implications for both theory and practice. At the theoretical level, it suggests that adequate arrangements should be interpreted as a bridge between corporate law, management control and knowledge management. The legal requirement to detect crisis becomes meaningful only when translated into routines for knowledge creation and use. This interpretation avoids two opposite reductions. The first is formalism, where adequacy is treated as the presence of documents, charts and procedures that do not affect decision-making. The second is technicism, where predictive software is treated as sufficient evidence of good governance. In both cases, the organisation may appear compliant while remaining unable to learn from weak signals.

The knowledge-management perspective also clarifies the relational dimension of crisis prevention. A firm is not an isolated calculator. It is embedded in financial, commercial, professional and community relationships. Early warning systems should therefore support dialogue among directors, managers, auditors, advisors, banks and other stakeholders. AI can improve the quality of such dialogue by producing more timely and structured information. Yet dialogue remains essential because recovery options depend on trust, credibility and negotiated commitments. In this sense, intelligent arrangements are also collaborative arrangements.

For boards of directors, the framework implies a shift in agenda. Directors should not limit their attention to annual accounts or retrospective reports. They should periodically examine forward-looking indicators, model assumptions, risk thresholds and the quality of managerial responses. They should ask whether the organisation can explain its alerts, whether different functions share the same interpretation of risk and whether corrective actions are implemented. The board's responsibility is not to become a data-science team, but to ensure that data science is governed as part of the control environment.

For control bodies and internal auditors, AI introduces the need to extend assurance activities to data and models. Traditional checks on accounting processes remain necessary, but they should be complemented by review of data lineage, access rights, model change management, validation frequency and exception handling. The audit question is no longer only whether a report is correct. It is also whether the system that produces predictions is reliable, explainable and monitored. This is especially important when AI outputs influence decisions on financing, restructuring, investments or communications with stakeholders.

For small and medium-sized enterprises, the framework should not be interpreted as a call for expensive technological infrastructures. Proportionality remains essential. Many smaller firms can benefit from simpler dashboards, automated cash-flow monitoring, structured checklists and external advisory support. The point is not the adoption of a particular technology, but the creation of a disciplined link between signals and decisions. AI may be delivered through accounting platforms, banking tools or advisory services, but the entrepreneur and governance bodies must still understand the logic, limitations and consequences of the information received.

The discussion also highlights a broader implication for intelligent and sustainable organisations. AI-enabled crisis governance is not only defensive. It can strengthen innovation and resilience because it improves the organisation's capacity to learn from data, coordinate responses and allocate resources. A firm that understands its vulnerabilities can invest more selectively, protect critical capabilities and design more robust relationships

with stakeholders. In this respect, crisis prediction and sustainable innovation are connected. Both require the conversion of dispersed knowledge into collective action under uncertainty.

7 Conclusions

This paper has argued that the integration of AI into adequate organisational, administrative and accounting arrangements represents a significant opportunity for sustainable crisis governance, but only under precise organisational conditions. The Crisis and Insolvency Code has made prevention a central duty of corporate governance. AI can support this duty by improving the detection of weak signals, integrating heterogeneous data and generating forward-looking scenarios. Nevertheless, AI does not remove the need for human judgement. On the contrary, it makes the quality of judgement more visible because directors must interpret, challenge and document the meaning of algorithmic outputs.

The main contribution of the paper is the conceptualisation of adequate arrangements as intelligent knowledge infrastructures. This view connects legal compliance with organisational learning and sustainable value creation. Adequacy is not a static state. It is a dynamic capability through which the firm senses emerging risk, interprets it collectively and acts before value is irreversibly destroyed. AI can enhance this capability when it is governed through reliable data, transparent models, clear responsibilities and ethical safeguards.

The paper is limited by its conceptual nature. Future research could test the proposed framework through case studies of firms that have adopted AI-based monitoring systems, comparative analysis across sectors, or empirical evaluation of how boards use predictive information. Further research is also needed on the role of professional advisors, banks and public institutions in creating shared early warning ecosystems. A promising direction concerns the relationship between crisis prediction and sustainability indicators, because environmental and social vulnerabilities may become financially material before they appear in traditional accounting measures.

For practitioners, the message is direct. The value of AI in crisis governance depends less on technological novelty than on organisational discipline. Firms should begin by mapping information flows, clarifying responsibilities, improving data quality and defining escalation routines. Only then can predictive systems become meaningful tools for continuity and resilience. The intelligent organisation of the future will not be the one that delegates governance to algorithms, but the one that combines human, collective and artificial intelligence in accountable decisions for sustainable value creation.

8 Acknowledgements

The authors acknowledge the relevance of the International Forum on Knowledge Asset Dynamics (IFKAD) debate on intelligent knowledge for sustainable organisations as the scholarly context within which this paper is positioned.

9 Ethics declaration

This paper is conceptual and does not report empirical research involving human participants, personal data, organisations observed in the field or experimental procedures. Ethical clearance was therefore not required.

10 AI declaration

Generative AI assistance was used to support language drafting, editorial structuring and consistency checks during the preparation of this manuscript. The conceptual framing, selection of sources, interpretation of the literature and final responsibility for the submitted text remain with the authors.

References

Airoldi, G., Brunetti, G. and Coda, V. (1994) Economia aziendale, Il Mulino, Bologna.

Altman, E.I. (1968) 'Financial ratios, discriminant analysis and the prediction of corporate bankruptcy', The Journal of Finance, Vol. 23, No. 4, pp 589-609.

Andrei, P. (2021) 'Adeguati assetti e responsabilita degli amministratori', Rivista dei Dottori Commercialisti, Vol. 72, No. 4, pp 645-670.

Argote, L. and Ingram, P. (2000) 'Knowledge transfer: a basis for competitive advantage in firms', Organizational Behavior and Human Decision Processes, Vol. 82, No. 1, pp 150-169.

Assonime (2020) La gestione del rischio e gli adeguati assetti nel Codice della crisi, Assonime, Rome.

Balcaen, S. and Ooghe, H. (2006) '35 years of studies on business failure: an overview of the classic statistical methodologies and their related problems', The British Accounting Review, Vol. 38, No. 1, pp 63-93.

Banca d'Italia (2019) Linee guida sul sistema dei controlli interni, Banca d'Italia, Rome.
Barboza, F., Kimura, H. and Altman, E. (2017) 'Machine learning models and bankruptcy prediction', Expert Systems with Applications, Vol. 83, pp 405-417.
Bauco, C., De Luca, R., Lucido, N. and Pagani, A. (2023) Assetti organizzativi, amministrativi e contabili: profili civilistici e aziendalistici, Consiglio Nazionale dei Dottori Commercialisti e degli Esperti Contabili, Rome.
Beaver, W.H. (1966) 'Financial ratios as predictors of failure', Journal of Accounting Research, Vol. 4, pp 71-111.
Borgonovi, E. and Anessi Pessina, E. (2018) Economia delle aziende, Egea, Milan.
CNDCEC (2020) Linee guida per la valutazione dell'adeguatezza degli assetti organizzativi, amministrativi e contabili, CNDCEC, Rome.
CNDCEC (2021) Norme di comportamento del collegio sindacale delle societa non quotate, CNDCEC, Rome.
CNDCEC and FNC (2023) Assetti organizzativi, amministrativi e contabili: check-list operative, Consiglio Nazionale dei Dottori Commercialisti e degli Esperti Contabili and Fondazione Nazionale di Ricerca dei Commercialisti, Rome.
Coda, V. (2010) Governare l'impresa, Il Mulino, Bologna.
Cortese, D. (2020) Assetti organizzativi e governo dell'impresa, Giappichelli, Turin.
Davenport, T.H. and Prusak, L. (1998) Working Knowledge: How Organizations Manage What They Know, Harvard Business School Press, Boston.
D.Lgs. 12 gennaio 2019, n. 14, Codice della crisi d'impresa e dell'insolvenza.
D.Lgs. 13 settembre 2024, n. 136, Disposizioni integrative e correttive al codice della crisi d'impresa e dell'insolvenza di cui al decreto legislativo del 12 gennaio 2019, n. 14.
European Parliament and Council (2024) Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence, Official Journal of the European Union, 12 July.
Floridi, L. and Cowls, J. (2019) 'A unified framework of five principles for AI in society', Harvard Data Science Review, Vol. 1, No. 1.
Fortunato, G. (2020) La gestione anticipata della crisi d'impresa, Giappichelli, Turin.
Grant, R.M. (1996) 'Toward a knowledge-based theory of the firm', Strategic Management Journal, Vol. 17, Winter Special Issue, pp 109-122.
Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F. and Pedreschi, D. (2018) 'A survey of methods for explaining black box models', ACM Computing Surveys, Vol. 51, No. 5, Article 93.
Hahn, T., Pinkse, J., Preuss, L. and Figge, F. (2010) 'Tensions in corporate sustainability: towards an integrative framework', Journal of Business Ethics, Vol. 99, No. 2, pp 217-235.
Jarrahi, M.H. (2018) 'Artificial intelligence and the future of work: human-AI symbiosis in organizational decision making', Business Horizons, Vol. 61, No. 4, pp 577-586.
Lessmann, S., Baesens, B., Seow, H.V. and Thomas, L.C. (2015) 'Benchmarking state-of-the-art classification algorithms for credit scoring: an update of research', European Journal of Operational Research, Vol. 247, No. 1, pp 124-136.
Marchi, L. (2011) Sistemi di controllo direzionale, Giappichelli, Turin.
Ministero della Giustizia (2022) Linee guida sulla composizione negoziata della crisi, Ministero della Giustizia, Rome.
Nonaka, I. and Takeuchi, H. (1995) The Knowledge-Creating Company, Oxford University Press, Oxford.
Ohlson, J.A. (1980) 'Financial ratios and the probabilistic prediction of bankruptcy', Journal of Accounting Research, Vol. 18, No. 1, pp 109-131.
Onida, P. (1971) Economia d'azienda, UTET, Turin.
Pisoni, P. (2022) 'Assetti organizzativi e flussi informativi nella governance', Contabilita e Finanza, Vol. 3, pp 45-62.
Quagli, A. (2019) La continuita aziendale nel Codice della crisi, Giappichelli, Turin.
Raisch, S. and Krakowski, S. (2021) 'Artificial intelligence and management: the automation-augmentation paradox', Academy of Management Review, Vol. 46, No. 1, pp 192-210.
Ribeiro, M.T., Singh, S. and Guestrin, C. (2016) 'Why should I trust you? Explaining the predictions of any classifier', Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp 1135-1144.
Shumway, T. (2001) 'Forecasting bankruptcy more accurately: a simple hazard model', The Journal of Business, Vol. 74, No. 1, pp 101-124.
Teece, D.J. (2007) 'Explicating dynamic capabilities: the nature and microfoundations of sustainable enterprise performance', Strategic Management Journal, Vol. 28, No. 13, pp 1319-1350.
Unioncamere (2022) Manuale operativo per la composizione negoziata, Unioncamere, Rome.
Van der Byl, C.A. and Slawinski, N. (2015) 'Embracing tensions in corporate sustainability: a review of research from win-wins and trade-offs to paradoxes and beyond', Organization & Environment, Vol. 28, No. 1, pp 54-79.
Wilson, R.L. and Sharda, R. (1994) 'Bankruptcy prediction using neural networks', Decision Support Systems, Vol. 11, No. 5, pp 545-557.

Leadership and Organizational Development in Remote Working Settings: Challenges, Opportunities and Empirical Evidences

Antonio Lerro[1], Daniela Carlucci[1], Francesco Santarsiero[1], Giovanni Schiuma[2]
[1] University of Basilicata, Potenza, Italy
[2] LUM University, Casamassima (BA), Italy

Abstract

In the last years it has been widely recognized the great relevance to identify and manage leadership traits in agile and remote working organizational settings. Entrepreneurs and top managers are increasingly playing a fundamental role in addressing the organizational changes determined by digital transformation and the adoption of new ways of working by rethinking consolidated structures and processes, fostering a culture of innovation and leading HR to adopt new mindsets, knowledge, and attitudes to thrive under continuous change and challenges. Despite this importance and the attention recently paid by academic and managerial studies trying to shed light on these issues, a full and shared setting of findings, insights and managerial and policy guidelines about the new styles of leadership and the required characteristics for a leader best suited the creation and the management of agile and remote working teams is still limited and also results are presented often in a jeopardized way. This paper analysed the main results of a scoping literature review on the main trends regarding challenges and opportunities for leadership in agile and remote working and the characteristics and skills required to entrepreneurs and managers to lead the organization in remote working contexts. The theoretical insights have been then examined and tested in Basilicata, a rural region in the southern Italy. In such a prospect, the research explored whether and how companies headquartered in Basilicata adopt remote working practices and analysed the leadership traits essential for their successful implementation.

Keywords: Leadership; Organizational Development; Remote Working; Rural Areas.

Paper type: Academic Research Paper

1 Research Background

Different scholars have recently argued that the exponential development and adoption of increasingly high-performing technologies and the significant impact of the digital transformation on the business processes have determined – among the different effects – also the emergence of different ways of working characterised by the possibility to perform at any location and at any time (Choudhury et al., 2021). While organisations of different industries were getting for the technology-driven future, the world has been unexpectedly hit by the COVID-19 pandemic. Beyond the dramatic aspects of the humanitarian crisis, it determined a breakthrough of all the economic activities. Accordingly, in order to guarantee economic survival and business continuity avoiding social contacts, companies and organisations encouraged and supported employees to work from home as much as possible. In this way, the most significant organisation and workplace design shock of our recent lifetimes was happening (Contreras et al., 2020; Hoff, 2021). However, it is important to underline that, before the COVID-19, these ways of working were steadily growing globally across many industries and infrastructures already existed (He et al., 2020): in this sense, the pandemic only accelerated this processes (Varma et al. 2022).

Consequently, the expressions *"agile working"* and *"remote working"* , as well as the related *"smart working"* and *"teleworking"*, emerged as mainstream concepts to identify and synthesize these new trends as well as the technological, organizational and human resource-related variables involved, the positive and negative consequences and finally the effects, the opportunities and the challenges both at individual, organizational and societal level (Grant, 2020). Accordingly, a full comprehension of these new approaches and related implications for HRM and a suitable and coherent leadership capable to serve as a guide for organizations and for existing and future leaders is fundamental for scientists as well as for practitioners involved on these fields.

The notion of "agility" is traditionally related to the organizational capacity to identify new trends, opportunities and threats, to be fast, pro-active and resilient, to have change-ready mindset to keep moving, changing and adapting to the general and task environment in which operates or could operate, mainly innovating business modelling, rethinking products and services to be addressed to actual and potential customers/users, revisiting operations, exploring the value of new resources and assets and defining a new platform of commercial and non-commercial relationships (Zheltoukhova, 2014). In relation to work practices, different scholars have focused the perspective of the *"agile working"* and have referred to agility as *"the ability to react quickly to fragmented world market situations full of constant and unexpected changes"* (Goldman,

1995), as the set of practices used to determine solutions using cross-functional and self-organized teams (Denning, 2016), as the *"summa of the adaptability, the quick response and the great execution"* (Accenture, 2015), as the capacity of a company to predict, perceive and respond to market volatility in a way that creates value (Balog, 2020). Other scholars, instead, have paid attention to the perspective of the "workforce agility", defined as the capacity of the employees to be *"well-trained and flexible, capable to adapt quickly and easily to new opportunities and market circumstances"* (Muduli, 2013). The same author also identified seven key attributes of an agile workforce, such as adaptability, flexibility, positive attitude towards development, speed, collaboration, competence and information (Richter et al., 2018). Shafer et al. (2001) define the agile workforce as a workforce each of whom is trained to do multiple jobs in the company to which they belong, while Nawaz and Gomes (2020), in connection to this, underline the necessity of replacing multi-level and rigid management structures with leaner and more flexible forms to better receive feedback allowing adaptation to customers requests and needs. The effect of the adoption of agile forms on organizational HRM practices is also a topic of the Manifesto for the Agile Development of Human Resources (2017) stating that *"the new mission of HRM is to push people towards new frontiers by motivating them and exploiting their strengths so that required goals can be achieved"*. This implies that HR managers should encourage the growth of individual talents by making them responsible for the role they play in the company, also in relation to the greater freedom in organizing and planning their activities anywhere and at any time (Pontillo et al., 2022).

Together with the agile working, a further relevant concept introduced to analyse the new work environments and practices is remote working, often used as umbrella notion embracing other terms such as smart working (Bednar and Welch, 2020; Boorsma and Mitchell, 2011; Lake, 2013; McEwan, 2016), teleworking, working at home, and home-based work (The International Labor Organizations, 2020; Nayani et al., 2018). It can be considered, then, a wide-ranging concept covering any paid work performed from a distance in any place different from the physical presence in the organization where employees meet organizational objectives through ICTs at their preferred time (Golden et al., 2008; Lautsch et al., 2009; Wojcak et al., 2016).

Despite different contributions have been provided to define and operationalize it (Mirchandani, 2020; Sullivan, 2003), unfortunately remote working till now does not have a universally-accepted definition and often it is overlapped with the agile working one (Tursunbayeva et al., 2021; Pontillo et al., 2021; International Labor Organization, 2020). This orientation is confirmed also by Gastaldi *et al.* (2014) describing remote working as *"an agile and dynamic way of working that leads to high performance, greater productivity and better satisfaction" for organizations and employees".* This tendency to use the concepts interchangeably is confirmed by the analysis of the research aimed to explore the nature, the effects and the advantages and disadvantages determined by such ways of working as well as to the challenges and opportunities that will be analysed in the following sections. For simplicity purposes, in the next sections the two concepts will be used merging them in the unique wider notion of remote working.

2 Challenges and Opportunities in Remote Working Settings

A rather wide literature has highlighted, in the last years, the advantages and the risks of the remote working (Contreras et al., 2020; Varma et al., 2022; Wojcak et al., 2016; Steude, 2021). Understanding and deeply analyzing these aspects represent a fundamental propaedeutic conditions to try to identify and delineate challenges and opportunities for leadership. In order to carry on a rigorous analysis of these multifaceted issues, following the main framing used by Steude (2021), it could useful to distinguish the perspectives of analysis, considering separately the *employees' view*, the *company view* and *the social view*. In fact, although very interconnected, the same phenomena may determine different perceptions and usefulness.

According to an *employees' view*, in terms of advantages, well-being is identified as the main positive effect (Charalampous, 2020; Hill et al., 2008), thanks to the greater flexibility and a better balance between private and professional life enabling higher level of productivity and efficiency (Gastaldi et al., 2014; Zheltoukhova, 2014; McEwan, 2016; Allen et al., 2015; Martin and MacDonnel, 2012). On these basis, empirical studies have found favorable outcomes of agile and remote working in specific elements such job performance, job satisfaction, work-life balance (Fedakova and Istonova, 2017), reduced rates of stress and better productivity (Gaiendran and Harrison, 2007; Pavlova, 2019), time of commuting, costs savings related to daily transport, reduced turnover intentions (Kossek et al., 2006; Fonner and Roloff, 2010; Coenen and Kok, 2014; Vega et al., 2015; Otham et al., 2009) till to reach a better employees' quality of life and happiness (Azarbouyeh and Naini, 2014; Kazakami, 2020). Always in this vein, Bentley (2014) highlights effectively the role of the technological support in activating these positive dynamics, while Morgan (2004) underlines the possibility to increase job opportunities for women and people with disabilities. However, also negative consequences have been mapped, such as isolation, loneliness and burnout (Moss, 2018) and forms of technostress related to multitasking,

permanent connectivity, anxiety about the possible shrinking of career prospects owing to reduced visibility (Maruyama and Tietze, 2012), overload of information (Kraut, 1989; Moore, 2006) capable to determine increasing demotivation (Wojcak et al., 2016; Fedakova and Istonova, 2017), work-family struggles (Golden et al., 2008) till to quite-quitting behaviours and workforce turnover.

According to a *company view*, advantages are traced in contributing to the organizational commitment (Nakrošien˙et al., 2019; Klopotek, 2017), in reducing the rate of absenteeism, in enhancing the process of standardization of procedures (Cortellazzo et al., 2019), in influencing the reputation and the corporate image due to the "*green-oriented*" behaviours related to reduction of traffic and daily transports, pollution, and use of energy sources within the workplaces (Giovanis, 2018). Moreover, companies may have access to specialized expertise, regardless of the team members' location, which allows companies to find more creative solutions (Malhotra et al., 2007). At the same time, different scholars (Contreras et al.,, 2020; Steude, 2021; Varma et al., 2022) highlight that also potential disadvantages emerge, as negative impact on the organizational culture and values (Wojcak et al., 2016; Ware and Grantham, 2010; Flood, 2019), difficulties in managing conflicts, team building, team working and engagement (Cooper and Kurland, 2002; Jaiswal and Arun, 2020; Younger and Kearns, 2017), difficulties of monitoring and coaching remote workers and preserving experiential knowledge (Golden et al., 2008), possibilities of decreasing mutual trust and company loyalty due to the physical absence from the company (Nayani et al., 2018), risks for the IT security of the company (Steude, 2021).

Finally, according to a *social view*, reduction of pollution and reduction of time for transportation and commuting are identified as the main environmental and social benefits. Also upgrading of the interest and the attractive capacity of the rural areas and the reduction of the cities' congestion can be considered further elements of advantages (Steude, 2021). At the same time, exacerbation of the digital divide between countries, and above all between organizations and employees accessing and using these opportunities and traditional and not-ready ones can be identified as further dysfunctionalities. On these basis, in the last times, it strongly emerges the need to define and analyse which kind of leadership styles are best suited and fitted to the agile and increasingly digitised organizational contexts. In particular, the profiling of the remote leader is fundamental to lead modern organisations to identify and select effectively managers capable to successfully meet the challenges and the opportunities of the new modelling of the work activities.

3 Leadership in Remote Work Settings: Key Characteristics and Competences

Identifying the traits and competences needed for leadership in remote work settings is crucial for generating suggestions for organizations adopting these modes of working and uncovers some potential needs of leaders who might face the challenges and opportunities of "digital revolution". Recently, different scholars have explored the topic of leadership in the context of remote work, motivated also by the spread of the digital transformation process in organizations and the effects produced by the pandemic (Avolio et al., 2014; Van Wart et al., 2019; Coun et al., 2021; Terkamo-Moisio et al., 2022; Harris and DeFlaminis, 2016; Iannotta et al., 2020; Tandon, 2021; Contreras et al., 2020; Cortellazzo et al., 2019; Pontillo et al., 2022).

In their work Zeuge et al. (2020) identify several emerging key leadership characteristics in remote work settings. Drawing on the results of a narrative literature review on leading virtual teams, the scholars point out which behaviours and traits a good virtual leader should have. According to the scholars a leader operating in remote working "should build trust, be empathetic and be open to diverse groups (starting with the tolerance for several time zones). At the same time, it is her or his responsibility to create a culture of "belonging" and 'being there for one another', 'caring', 'listening' and empathy. What is required here is the ability to communicate and to have emotional intelligence. A virtual leader is always available, approachable, addressable, and open. She or he demands by promoting an open mindset, because she or he is a good example herself or himself" (p. 7). Additionally, Zeuge et al. (2020) underscores that leading team in remote working requires building trust and a culture of collaboration and support among team members. According the authors, leaders should manage conflicts and identify frustration in the team early on as well as organize regular meetings, either formal or informal, where team members can socialize.

The crucial importance of meeting all remote workers, also in person if possible, and enhancing social interactions to reduce misunderstanding and feelings of loneliness typically experienced by remote workers, have been addressed by several authors (e.g. Brake, 2006; De Rosa, 2009; Thomas and Bostrom, 2008). Other authors have analysed the style to build effective interactions and relationships in remote working. According to some authors (e.g. Hart and McLeod, 2003; Madlock, 2018), remote leaders should adopt a task oriented leadership style to build effective and trusted relationships and job satisfaction in teams. They should promote among workers an intensive task-related orientation at a distance, then facilitate face-to-face social opportunities, and finally allow workers return to remote locations to continue work. There are some empirical

evidences that a task-oriented leadership style would produce more communication satisfaction, job satisfaction and telecommuter organisational commitment than relational-oriented leadership. However, these findings contrast with other studies (e.g., Brake, 2006; Hunsaker and Hunsaker, 2008; Maduka et al., 2018 pointing out the relevance of relationship-oriented leadership. According to these studies, leaders should be relationships-oriented as stronger relationships support a more developed mutual understanding, because the messages exchanged are more frequent and briefer and characterized by a more proactive focus on the problems or challenges faced by the work. This is particularly important considering the fact that challenges and conflicts are increasingly difficult to resolve in remote communities compared to traditional contexts (Brake, 2006). In remote setting increased technological frustrations, barriers to performance, loss of interest, and cultural differences can undermine problem solving process. Remote leaders must adopt a proactive role and establish rules and procedures to resolve problems that arise (e.g., Harris et al., 2001).

The crucial importance of relationships, meetings, building trust, effective communication is widely highlighted also by Contreras et al. (2020). These latter scholars, investigating tele-working and e-leadership, underline that e-leaders must develop technical competencies to monitor, coordinate, and align work through technology-supported structures, reducing barriers. They must be expert with latest ICTs and create awareness for teleworking. Remote leaders should stay updated on technological advancements, seek out technological solutions, and remain informed about recent developments in their areas of expertise (e.g. Cortellazzo et al., 2019; Harris et al., 2001; Maduka et al., 2018).

As remote environments can influence workers' behaviors, emotions, thoughts, and performance, according to Contreras et al. (2020) e-leaders must reduce physical and psychological distances, by being close to employees, creating trust through interrelationship and avoiding to organize fragmented tasks. Moreover, they confirm the importance of shared leadership as stimuli to team members to take responsibility for their work and the project, promoting identification within the team and initiating action flows for goal achievement.

Analogously, Flood (2019) states that “beyond technical and business skills, leadership must acquire specialized aptitudes for coaching, adaptability, career guidance, communicating, interpersonal relationships, and developing cultural sensitivity” (p. 3), while other scholars highlight the fundamental role of the communication abilities (e.g. Roy, 2012; Zaccaro and Bader, 2003; Cortellazzo et al., 2019; Owens et al., 2011; Harris et al., 2001). Very interesting is also the work Pontillo et al. (2022). Through a scoping literature review of 54 relevant studies, they identified leadership qualities, potential skills, and how varied leadership approaches can foster fitting innovative work behaviors, in agile and remote work settings. They specifically identified styles of leadership as directive, catalyst, resonant, servant, transformational, simplifying, and distributed.

The review of management literature shows that to date the research on the leadership in remote settings is still limited, but also wide-ranging and promising. The researchers in their studies have examined one or more traits and competences characterizing leaders and their leading action in remote setting. The developed analysis of management literature on leaderships characteristics in remote working, suggest some key competences that leaders operating in such a context should have. They can be grouped as showed in Figure 1. Certainly the list of the reported distinguishing competences is not exhaustive and additionally may vary depending on the organization's culture, industry, and the nature of the remote work being conducted.

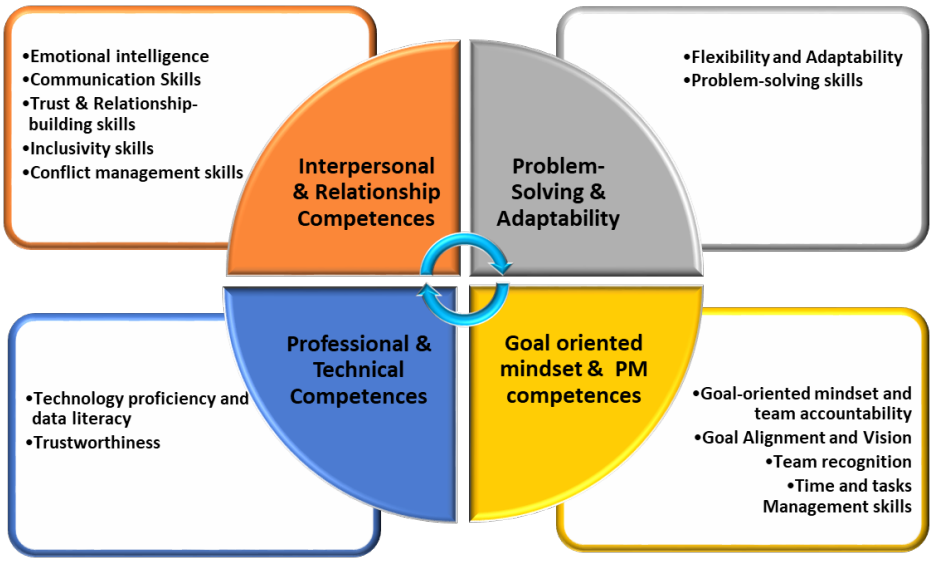


Figure 1. Leadership competences in remote working setting (our elaboration)

4 Empirical Analysis

The findings and insights from the desk analysis have been examined and tested in Basilicata, a region in southern Italy. Basilicata is a rural area with low population density, reliance on agriculture and services, mainly related to Public Administration bodies, predominance of natural landscapes, and the presence of small and rather isolated communities with unsatisfactory urban infrastructures and very limited cities acting as economic and social hubs for the whole region. Depopulation, particularly among young people, is a noteworthy phenomenon in Basilicata, driven by limited job opportunities and the migration of younger generations abroad in search of education and employment. Accordingly, remote working could have the potential to nourish socio and economic development of rural area, fostering economic growth, social revitalization, and environmental sustainability. In such a prospect, the research explored whether and how companies headquartered in Basilicata, or primarily operating within the region, adopt remote working practices and analysed the leadership traits essential for their successful implementation.

The empirical analysis was developed according to three main stages. The *first stage* was aimed at identifying the companies and the organizations to be potentially involved in the study. For operative purposes, it was divided into *three macro-activities*.

The *first macro-activity* involved a scoping desk analysis using different sources - mainly web-based - to collect preliminary inputs and suggestions regarding industries, companies and organizations based in Basilicata and applying remote working practices. Unfortunately, this activity revealed the limited use of remote working.

The *second macro-activity* was related to the contact and the engagement of experts and leaders of the local business and managerial world, and relevant actors and stakeholders such as Confindustria Basilicata and Federmanager Basilicata. These interactions confirmed the concerns about the identification of suitable companies and organizations to participate in the study due to the region's specific gaps, including a generally low availability and use of technological and digital tools, a very limited number of companies with robust organizational structures and an explicit need for remote working practices, and little attention to HRM, to name a few of the most commonly cited issues.

Matching the results and the insights derived from the first macro-activity with the very useful suggestions gathered in the second macro-activity through the interactions with the local experts and the stakeholders, it was possible to elaborate, as *third macro-activity*, consisting in delineating a preliminary list of companies and organizations based in or mainly operating in Basilicata that could potentially be considered and contacted for the ongoing empirical study.

In the *second stage* of the empirical investigation, the companies and the organizations identified in the preliminary list were contacted by mail and phone. All of them responded positively to the information asked and were open to engaging with the research team. Following this initial contact, the list was revised and refined to include only those companies that were structurally implementing remote working practices. As hypothesized, these companies were all operating in industries related to ICT products and services.

Finally, in the *third stage* of the empirical investigation, the five companies selected were contacted again to schedule specific dates, times, and modalities for submitting and sharing the questionnaire that had been developed for this purpose and to have direct interviews virtually or on-site, aimed to explore and capture additional thoughts and qualitative comments from the questionnaire's respondents. The questionnaire was structured into two main sections: the first section gathered important information for company profiling, while the second section focused on understanding the modalities and impact of remote working practices within the organizations, as well as the related leadership styles and practices adopted by entrepreneurs, CEOs, and top managers. A synthesis of the companies' profiles – labelled anonymously using greek letters for privacy purposes - and the global insights derived from the empirical investigation are summarized in the following.

The first company analysed was *Alfa s.r.l.*, a limited liability company based in Matera engaged in "*consultancy in the technology sector of information*". Founded in 2010, *Alfa* is a software house specialized in offering software products and services based on advanced technologies. Leveraging its accumulated expertise, the company has developed a deep know-how in all phases of the software life cycle, including design, maintenance, and support.

Founded in 2000, the second company analysed was *Beta*, which represents one of the main, and longest-standing, existing realities in the Lucanian ICT sector. Currently, the company is configured as a corporate structure of a joint stock company and has its main headquarter in Basilicata, in Sant'Angelo Le Fratte, near the city of Potenza. It operates through eight plants, whose five are based in Basilicata, and counts over 1.000 employees.

The third company identified as empirical setting was *Gamma*. It is one of the oldest companies in the IT sector in Basilicata. The company locates its headquarter in Tito, in the province of Potenza.

The fourth company under investigation was *Delta*. The company deals with the "production of software not connected to the edition". It is based in Matera and is configured in the corporate structure of a joint-stock company (S.P.A). The company represents a thriving Lucanian company operating in the ICT sector, since 2001. The foundation of the company falls within the framework of the ESA project (The European Space Agency) called COSMO-SkyMed (Constellation of small Satellites for the Mediterranean Basin Observation), with the role of algorithmic developer for the first dual system (military and civil) of earth observation radar satellites.

Finally, is was identified the *Theta* company. It is a limited liability company (s.r.l.) founded in 2001 in Matera, with operational offices also in Potenza, Bari and Milano. The company's core business is focused on the outsourced contact center sector, i.e. the delegation of call centre activities by a company to a specifically dedicated external organisation and market research.

5 Discussion and Implications

The study highlighted that remote working practices in the analyzed cases are mainly adopted by companies with a mature and established corporate history. These companies have experienced growth over the years in terms of both turnover and employee numbers. The COVID-19 pandemic, from this point of view, was often perceived by these companies as a catalyst for starting a transition process towards a qualitative leap in their organisation. This leap has included the dynamics linked to remote working practices.

Initially adopted as a mere mandatory measure for company operations, remote working has ended up becoming an essential pillar of business re-organization. However, the perception of its need and effectiveness - as well as its adoption - has not been uniformly positive. The spectrum of reactions varies from more typically "techno-enthusiastic" companies, that proactively embrace the benefits of remote work- compensating for the apparent loss of control over the employees and satisfied about the workforce results in terms of productivity, to organizations that are more reluctant to accept these working practices.

A common concern within all the organizations involved in the research is represented by the idea that it is considerably more complex, for employees joining the company from scratch, easily building their active participation in the corporate culture, remotely conveying those values and exemplary behaviours that progressively cement employees' adherence to a certain value system and professional mentality.

At the same time, a commonly perceived benefit is the possibility of recruiting professionals without geographical constraints. On the one hand, this is an element to increase an organisation's attractiveness and competitiveness within a territory, by allowing access to the specific know-how that is needed and minimizing the training required for employees who may lack certain skills. On the other hand, this dynamic also takes the form of a defensive and preventive measure against internal turnover and external dumping. As highlighted in the interviews, remote working has often allowed ICT organizations to minimize the theft of their human capital by big companies of international importance.

To counteract the challenges inherent in remote working, companies are trying to create alternative moments of conviviality for in-person interactions, sharing organizational culture, impressions, feelings and values. When asked about their chosen leadership style, based on the provided nomenclature, the answers were rather heterogeneous and debated: catalyst, servant and transformational - or a holistic sum of all the leadership styles presented – appear to be the driving methods in which the entrepreneurs and managers who participated in the survey best identified themselves. Table 2 summarizes the main benefits and challenges of adopting remote working perceived by the investigated companies.

Table 1. The main benefits and challenges of adopting remote working perceived by the analyzed companies

Benefits	Challenges
Increase of productivity and reduction of absenteeism: Higher workforce productivity, greater organizational commitment, and trust between employees and management, with a reduction in absenteeism.	**Difficulties in cultural absorption**: Greater challenges for new employees in absorbing the company's culture, values, and professional mindset.
Access to a broader talent pool: Recruitment is no longer limited by geographic location, allowing access to specialized skills and reducing the need for training.	**Challenges in preserving corporate culture**: Difficulties in transmitting and aligning the company's values, and challenges in consolidating a strong and shared corporate culture among remote workers
Costs saving: Reduction in costs related to decreased training needs, savings on expenses such as office-related costs, transportation, and energy.	**Complex onboarding**: Difficulties for new hires to become active participants in corporate culture and internal organizational dynamics.
Talent retention: Preventing turnover and protecting companies from losing staff to international competitors.	**Increased workers' isolation and loneliness & risk of techno-stress**

Benefits	Challenges
Improvement of work-life balance: Enhanced balance between employees' professional and personal lives, with reduced commuting times and greater scheduling flexibility.	**Management and monitoring of remote work**: Challenges in effectively monitoring and supporting remote employees, particularly concerning teamwork and conflict management
Improvement of the company's reputation	**Lack of face-to-face interactions:** The absence of physical interaction moments can lead to a reduced sense of belonging and engagement, with potential negative effects on team cohesion and collaboration
	IT security risks

Companies share both similar advantages and challenges of remote work. Although they have benefited in terms of improved productivity, recruitment, cost saving, reputation and employee well-being, they face several difficulties related to increased workers' isolation and loneliness, organizational culture absorption and preserving, IT security risks, and remote team management. To face these difficulties, the investigated companies have developed several initiatives often creating meetings, both virtual and in-person, to mitigate negative effects of remote working. In particular, there are several practices that the companies involved in the research are adopting to address the mentioned challenges. These are good practices that - despite the diversity of their business and organizational contexts - have common objectives, such as reducing isolation, strengthening corporate culture, and optimizing remote work performance assessment and management. Table 3 reports the good practices cited by the companies, clustered by the main objective.

Table 2. Good practices of the analyzed companies

Objective	Good practices implemented by the analyzed companies
Mitigating Isolation	· Introduction of "unexpected" elements to alleviate isolation and routine fatigue. · Creation of social events like "Start Days" and CUBE associations to foster connections. · Family-oriented social events to reduce isolation among remote workers. · Hybrid model tailored to employee proximity (remote for those farther away).
Strengthening Corporate Culture	· Use of hybrid interaction events to reinforce corporate culture.
Performance Assessment and Management	· Meritocratic performance evaluations based on quantitative and qualitative metrics. · Continuous training for team leaders to improve remote team management and engagement.

Regarding the traits characterizing an adequate leadership style referred to the adoption of remote working practices, empathy and communication skills emerge as fundamental for maintaining employee trust and engagement, collaboration and organizational cohesion. Certainly, these characteristics are challenging to cultivate and put into practice where daily face-to-face contact is lacking and is filtered through purely virtual forms of interaction. Looking ahead to the future of remote working, the process is perceived as inevitable. Conversely, it emerges a strong need to increase the compensation of the same by increasing opportunities for genuine human interaction.

In inner mountain areas, remote working can represent an important opportunity and has great potential to regenerate local economies and communities. However, it must be approached carefully. Good practices, such as creating hybrid remote working spaces, investing in continuous training for managers and entrepreneurs to acquire proper skills for managing remote teams and work, and nurturing a remote working culture, can promote long-term residency, attract expert workers from outside, reduce migration, preserve the survival of local businesses, and ensure their competitiveness. However, to avoid failure, it is crucial to consider local community needs, to avoid focusing merely on attracting remote workers from outside (e.g. digital nomads) and to balance the requests of local development with business growth. It is also important to have managers and entrepreneurs able to address the strategic and operational challenges of remote working. More generally, remote working can contribute to the growth of inner mountain economies and societies only if there is a balance between the opportunities of remote working, the need for sustainable development, and the integration of the local community with its needs and culture.

6 Acknowledgements

Funded by the European Union - NextGenerationEU, Project NODES – Nord Ovest Digitale E Sostenibile - Mission 4 Component 2 - ECS00000036 – Spoke 4, Flagship project 1, Research Module 4 Smart West.

7 Ethics Declaration

This research did not require formal ethical clearance, as it involved non-sensitive data collection through voluntary surveys and interviews with adult participants operating in a professional capacity. Informed consent was obtained from all participants, and anonymity and confidentiality were ensured throughout the study.

8 AI Declaration

No generative AI tools were used to produce, analyze, or revise the content of this paper. All data analysis, literature review, and writing processes were conducted by the authors using conventional academic methods.

Main references (complete list of references available on request)

Al-Malki, M., Juan, W. (2018) "Leadership Styles and Job Performance: a Literature Review", *Journal of International Business Research and Marketing*, Vol. 3, Issue 3, pp. 40-49.

Ardi, A., Djati, S. P., Bernarto, I., Sudibjo, N., Yulianeu, A., Nanda, H. A., and Nanda, K. A., (2020) "The relationship between digital transformational leadership styles and knowledge-based empowering interaction for increasing organisational innovativeness", *International Journal of Innovation*, Creativity and Change, Vol. 11, No. 3, pp. 259-277.

Bednar, P.M., Welch, C. (2020), "Socio-Technical Perspectives on Smart Working: Creating Meaningful and Sustainable Systems", *Information Systems Frontiers*, Vol. 22, pp. 281-298.

Charalampous, M., Grant, C.A., Tramontano, C., Michailidis, E. (2020) "Systematically reviewing remote e-workers' well-being at work: a multidimensional approach", *European Journal of Work and Organizational Psychology*, Vol. 28(1), pp. 51-73.

Choudhury, P., Foroughi, C., Larson, B. (2021) "Work-from,-anywhere: The Productivity Effects of Geographic Flexibility", *Strategic Management Journal*, Vol. 42(4), pp. 655-683.

Contreras, F., Baykal, E., Abid, G. (2020) "E-leadership and teleworking in times of COVID-19 and beyond: What we Know and Where Do We Go", *Frontiers in Psychology*, Vol. 11, pp.1-11.

Cortellazzo, L., Bruni, E., & Zampieri, R. (2019) "The role of leadership in a digitalized world: A review", *Frontiers in Psychology*, Vol. 10, 1938.

Coun, M.J.H., Edelbroek, R., Peters, P., and Blomme, R.J. (2021) "Leading Innovative Work-Behavior in Times of COVID-19: Relationship Between Leadership Style, Innovative Work-Behavior, Work-Related Flow, and IT-Enabled Presence Awareness During the First and Second Wave of the COVID-19 Pandemic", *Frontiers in Psychology*, Vol. 12, pp. 1–16.

Denning, S. (2016) "How to Make the Whole Organization 'Agile'", *Strategy & Leadership*, Vol. 44 (4), pp. 10–17.

Dinibutun, S.R. (2020) "Leadership: A Comprehensive Review of Literature, Research and Theoretical Framework", *Journal of Economics and Business*, Vol. 3(1), pp. 44-64.

Grant, C. (2020) Concepts, Terms and Measurement in Agile Working, in C. Grant and E. Russell (eds), Agile Working and Well-Being in the Digital Age, pp. 19–32, Cham: Palgrave Macmillan.

Harris, A., and DeFlaminis, J. (2016), "Distributed Leadership in Practice: Evidence, Misconceptions and Possibilities", Journal of Management Education, 30 (4): 141–146.

Iannotta, M., Meret, C., and Marchetti, G. (2020) "Defining Leadership in Smart Working Contexts: A Concept Synthesis", Frontiers in Psychology, Vol. 11, pp. 1–11.

Khan, Z.A., Nawaz, A., Khan, I. (2016) "Leadership Theory and Styles: A Literature Review", *Journal of Resources Development and Management*, Vol. 16, pp.1-7.

Klein, M., (2020) "Leadership Characteristics in the Era of Digital Transformation", *Business & Management Studies: An International Journal*, Vol. 8, No. 1, pp. 883–902.

Madlock, P. E. (2018) "The influence of leadership style on telecommuters in the insurance industry: a contingency theory approach.", *Journal of Leadership, Accountability & Ethics,* Vol. 15(2).

Maduka, N. S., Edwards, H., Greenwood, D., Osborne, A., & Babatunde, S. O. (2018) "Analysis of competencies for effective virtual team leadership in building successful organisations", *Benchmarking: An International Journal*, Vol. 25(2), pp. 696-712.

Manifesto for Agile HR Development. Available online at https://www.agilehrmanifesto.org/ (last accessed: November 24, 2021).

Moss, J. (2018), "Helping Remote Workers Avoid Loneliness and Burnout", *Harvard Business Review.* Available online at https://hbr.org/2018/11/helping-remote-workers-avoid-loneliness-and-burnout

Pontillo, S., Di Lauro, S., & Antonelli, G. (2022) "Defining the Leader in an Agile and Remote Working Environment", *PuntOrg International Journal*, Vol. 7(2), pp. 160–215, ISSN:2499-1333

Santarsiero, F., Lerro, A., Carlucci, D., and Schiuma, G., (2021) "Modelling and managing innovation lab as catalyst of digital transformation: theoretical and empirical evidence", *Measuring Business Excellence*, Vol. 26, No. 1, pp. 81-92.

Steude, D.H. (2021) "Challenges of Remote Leadership in a Digitalized Working World 4.0", *Sciendo*, doi.org/10.1515/mosr-2021-0005

Terkamo-Moisio, A., Karki, S., Kangasniemi, M., Lammintakanen, J., and Häggman-Laitila, A. (2022), "Towards Remote Leadership in Health Care: Lessons Learned from an Integrative Review", *Journal of Advanced Nursing*, Vol. 78 (3), pp. 595–608.

Van Wart, M., Roman, A., Wang, X., & Liu, C. (2019), "Operationalizing the definition of e-leadership: identifying the elements of e-leadership", International review of administrative sciences, Vol. 85(1), pp. 80-97.

Varma, A., Jaiswai, A., Pereira, V., Kumar, Y.L.N. (2022) "Leader-member exchange in the age of remote working", *Human Resource Development International*, Vol. 25(2), pp. 219-230.

When Does AI Become Knowledge? Organizational Conditions for Epistemic Learning in Public Administration

Emilia Romeo, Nicola Capolupo
San Raffaele Roma University, Rome, Italy

Abstract

The growing adoption of Artificial Intelligence (AI) in public administration has largely been framed in terms of efficiency gains and decision-making enhancement. However, this perspective overlooks a deeper transformation concerning the reconfiguration of organizational knowledge processes. This study aims to examine how AI contributes to knowledge creation, interpretation, integration, and institutionalization within public administrations, and to identify the organizational and institutional conditions that enable or constrain these processes. By viewing AI not just as a technical tool but as a system that shapes how knowledge is created, interpreted, and used, this study aims to provide a deeper understanding of public organizations as knowledge-based systems.

Methodologically, the research adopts a qualitative meta-synthesis of interdisciplinary literature on AI in public administration. A structured search was conducted in the Scopus database, resulting in a corpus of 108 relevant studies. The analysis combines computational text analysis techniques—such as lexical mapping and heuristic scoring—with an interpretive thematic synthesis. This approach enables the systematic identification of recurring knowledge processes, enabling mechanisms, and contextual constraints shaping the role of AI in organizational knowledge dynamics.

The findings reveal that AI operates as a knowledge-enabling and knowledge-processing infrastructure through three interrelated processes. First, AI enhances knowledge creation by identifying patterns and transforming data into actionable insights. Second, it supports knowledge interpretation and integration by enabling interactive and dialogical processes between human judgment and algorithmic outputs. Third, AI facilitates the institutionalization of knowledge through feedback loops, continuous learning, and adaptive system updates. However, the effectiveness of these processes depends on key organizational conditions, including transparent and participatory governance, experimental leadership, supportive organizational cultures, and adequate technical and infrastructural capabilities. Conversely, rigid hierarchies and control-oriented approaches tend to limit AI's role to procedural functions, constraining knowledge development.

From a theoretical perspective, the study contributes to knowledge management and public administration literature by integrating fragmented research streams into a unified framework that conceptualizes AI as a socio-technical knowledge infrastructure. It challenges traditional views of rational decision-making by highlighting the emergence of hybrid knowledge processes that combine human and algorithmic reasoning. Practically, the findings suggest that public organizations should design governance mechanisms that promote transparency, participation, and learning, invest in leadership and data literacy, and adopt flexible regulatory frameworks to fully leverage AI as a generative knowledge resource.

Keywords: Artificial Intelligence; Public Administration; Organizational Learning; Knowledge

Paper type: Academic Research Paper

1 Introduction

The advent of Artificial Intelligence (AI) in public sector has brought widespread optimism regarding its power to improve decision-making (Oryshchuk and Minenko, 2024). Nevertheless, this single exploitable view tends to overlook a more profound transformation that AI is affecting organizations, namely a reconfiguration of organizational knowledge processes, including knowledge creation, integration, and application in collective decision-making (Bolton et al., 2021; Leoni et al., 2024). Sticking with this perspective, public administrations can increasingly be understood as knowledge-based organizations, that is, arrangements wherein the functioning depend on the continuous production, sharing, and utilization of knowledge assets (Knorr Cetina, 2007; Patriotta, 2004). In these types of organizations, legitimacy and authority are not simply rooted in hierarchical domination but also in the capacity to generate, mobilize, and govern knowledge effectively. Therefore, the emergence of AI not merely alters bureaucratic practice but also transforms the underlying knowledge management mechanisms through which organizations identify problems, interpret information, and support decision-making processes. AI systems participate in organizational knowledge processes through shaping the manner in which evidence is collected, patterns are discerned, and interpretations are stabilized, thus contributing to the development and redistribution of knowledge across human and non-human actors.

While several studies have been done on AI adoption by the lens of governance, ethics, and automation (Camilleri, 2024), attempts at how it has been instrumental in altering knowledge creation, sharing, and utilization processes within public organizations under conditions of uncertainty (Albous and Alboloushi, 2025)

have been less emphasized. That is, AI is studied primarily as a technical innovation, while AI as a knowledge infrastructure enabling the development, integration, and application of organizational knowledge is seldom researched (Larsen, 2021). This perspective is particularly relevant when public administrations face complex challenges requiring not only faster decisions but also more effective knowledge integration processes capable of combining human judgment and algorithmic inference.

Thus, this research seeks to examine how AI is employed in the support and enhancement of organizational knowledge processes, including knowledge generation, interpretation, and integration within administration and under what organizational and institutional conditions such applications are feasible or constrained. This view challenges the technical-bureaucratic control dichotomy by examining how AI interacts with the knowledge management dimensions of public decision-making and thereby contributes to the construction of public administration as a dynamic knowledge-based organization in which knowledge itself is the critical site for governance.

2 Methodology

A qualitative meta-synthesis of current scientific literature on the interface between public administration and AI was conducted (Walsh and Downe, 2005). This research allows for the results of various studies to be combined while unlocking conceptual and analytical elements related to organizational knowledge processes and knowledge management practices in the public sector in the context of AI. The study employed a multi-step approach for qualitative meta-synthesis following Noah (2017) and Weinberger (2022).

Some details of the protocol employed are provided as follows. Two research questions were formulated to delineate the boundaries of knowledge processes and knowledge management dynamics of the study:

- *RQ1. What mechanisms facilitate the creation, integration, and application of knowledge through AI in public administration?*
- *RQ2. Under which organizational and/or institutional conditions do these knowledge processes and knowledge management functions flourish or stagnate?*

To collect the relevant material a structured search in Scopus database was carried out using the following prompt query:

TITLE-ABS-KEY ((“artificial intelligence” OR AI OR “machine learning” OR “generative AI” OR chatbot* OR “algorithmic system*”)
AND (“knowledge sharing” OR “knowledge management” OR “organizational learning” OR “knowledge transfer” OR “knowledge creation” OR “epistemic organization*” OR “epistemic community” OR “knowledge governance”)
AND (“public administration” OR “public sector” OR “government organization*” OR “public institution*” OR “public policy” OR “e-government” OR “digital government”)).

Following the protocol, the review included peer-reviewed academic studies from 2003 to date. This temporal boundary was deliberately adopted to capture the evolution of the epistemic mechanisms through which different types of algorithmic systems have been progressively integrated into organizational decision-making processes, rather than providing a technically uniform definition of AI. The inclusion criteria required that selected studies explicitly address the application of AI or algorithmic systems within the public sector in relation to knowledge processes, knowledge management, decision support, discovery, or policy innovation, while purely technical contributions without organizational implications were excluded.

Adopting a longitudinal perspective allows the analysis to account for the transformation of AI’s role in organizations over time. Earlier contributions, largely associated with rule-based systems, portray AI as exerting a predominantly procedural and normative influence, supporting routine decision-making, reinforcing standardization, and extending bureaucratic rationality. Accordingly, the mechanisms emphasized in this phase mainly concern efficiency and decision consistency.

However, with the emergence of more advanced approaches—particularly from the mid-2010s onwards, a qualitative shift becomes evident. AI increasingly shapes not only decisions themselves but also their underlying premises, enabling pattern recognition, hypothesis generation, problem reframing, and sensemaking processes. This transition marks the growing relevance of the mechanisms central to this study, namely AI as a cognitive infrastructure, an interpretive device, and a driver of organizational learning.

By incorporating studies across this temporal span, the meta-synthesis does not treat the evidence as homogeneous; rather, it adopts a processual perspective that allows the study to trace how these epistemic mechanisms intensify and transform alongside technological developments, thereby strengthening the explanatory power of the resulting conceptual framework.

Figure 1 displays the initial corpus of contributions per year retrieved by the analysis and shows steady growth over time, with a particularly marked increase in the most recent period. After a slow start between 2007 and 2017, production increased significantly from 2017 onwards. The period 2021–2026 marks a substantial quantitative leap, more than doubling previous values.

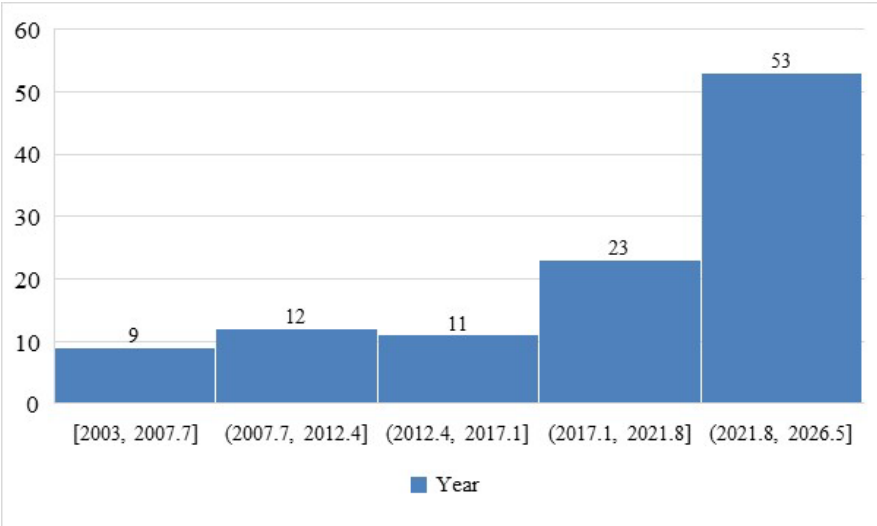


Figure 1: Pre-screening articles distribution per year

Table 1 shows a clear prevalence of conference-related contributions, with 43 conference papers and 16 conference reviews, representing over half of the total. These are followed by scientific articles (34), while book chapters (9) and books (4) are less frequent. Reviews and retracted articles are marginal, with only one case each.

Table 1**:** Types of contribution distribution

Type of contribution	Values
Article	34
Book	4
Book chapter	9
Conference paper	43
Conference review	16
Retracted	1
Review	1
Final value	108

To analyze the selected studies, a qualitative meta-synthesis supported by computational text analysis techniques was employed. The corpus of 108 articles was first examined through computational techniques such as lexical mapping and heuristic scoring, aimed at identifying recurrent concepts and patterns and supporting the systematic exploration of the literature. These techniques were used to enhance the transparency and consistency of the analytical process, rather than to replace interpretive analysis.

Subsequently, thematic synthesis was employed to integrate findings across different fields, including organizational studies, governance research, and information systems. This interpretive phase enabled the identification of recurrent knowledge processes, enabling mechanisms, and contextual constraints shaping the use of AI.

As such, through qualitative meta-synthesis, data were not simply aggregated but systematically interpreted to develop a conceptual framework that explains how AI reshapes organizational knowledge creation, integration, and application processes within public institutions.

3 Findings

The findings indicate that AI is a knowledge-enabling and knowledge-processing infrastructure, rather than a decision aid. Three interrelated knowledge processes emerge from the literature (RQ1).

First, machine-learning and data-driven systems enhance organizations' cognitive capacity knowledge creation capacity by allowing them to identify underlying correlations and patterns and converting information into potential hypotheses actionable knowledge and insights for public decision-making (Janssen et al., 2017).

Second, AI supports knowledge interpretation and integration in explainable and interactive systems that offer dialogue between human experience and algorithmic inference (Giraldi et al., 2024). These systems mediate interpretation rather than suppress it, enabling administrators to co-construct meaning, validate insights, and critically assess underlying assumptions embedded in the data.

Third, AI functions as an infrastructure for knowledge institutionalization and continuous knowledge development (Laihonen and Kokko, 2023). AI can power adaptive organizational knowledge processes through feedback loops, continuous model updates, and revisions based on errors.

Combined, such processes suggest that AI has the potential to transform public administration into a more exploratory, reflective, and adaptive knowledge-based system, rather than a purely rule-based mechanism (Babšek et al., 2025).

At the same time, the analysis concludes that the actualization of these cognitive capacities knowledge processes and knowledge management capabilities depends on a set of organizational and institutional conditions (RQ2). Governance arrangements characterized by transparency and participation are necessary to ensure that algorithmic systems operate as shared knowledge assets, rather than tools of control (Olsen et al., 2024).

Experimental leadership and deliberative practices contribute to the development of a knowledge-sharing culture, enabling the use of AI for knowledge exploration rather than mere surveillance or compliance. Furthermore, adaptable regulatory frameworks, adequate infrastructural resources, and critical capabilities—such as access to data, financial resources, and technical expertise—are required to support the effective development and use of AI-driven knowledge processes within public administrations.

Conversely, rigid hierarchies, technocratic mindsets, and excessive regulatory oversight tend to suppress knowledge creation and sharing processes, limiting the role of AI to procedural or predictive functions.

Therefore, the analysis shows that AI's contribution to knowledge management and organizational knowledge development in public administration emerges from an organizational transformation in which human judgment, knowledge processes, and algorithmic reasoning are effectively aligned (Mainardi, 2025; Tangi et al., 2025).

4 Discussion and Conclusion

This study contributes to the existing debate on AI and governance by reframing public administration as a knowledge-based system in which the value of AI extends beyond efficiency and automation to encompass its role in supporting organizational knowledge dynamics. Specifically, the findings highlight how AI enables the development and coordination of key processes related to the generation, interpretation, and use of knowledge within decision-making contexts (Janssen et al., 2017).

From this perspective, the contribution of the study lies in showing that the impact of AI is contingent upon the presence of appropriate organizational and institutional arrangements. In particular, governance architectures, leadership approaches, and accountability frameworks play a critical role in enabling AI to function as a shared and generative resource, rather than as a purely technical or control-oriented tool (Olsen et al., 2024).

Figure 1 illustrates the conceptual framework emerging from the analysis, positioning AI as a knowledge infrastructure embedded within organizational processes. The model integrates the study's key findings into a systemic representation that highlights how AI contributes to the generation, transformation, and institutionalization of knowledge in public administration.

At the core of the framework are three interrelated knowledge processes: knowledge creation, knowledge interpretation and integration, and knowledge institutionalization. These processes reflect the central mechanisms identified in literature. These processes are not conceived as linear stages but as mutually reinforcing, dynamically interconnected activities. This perspective aligns with the knowledge management literature, which emphasizes the continuous and iterative nature of knowledge development within organizations (Nonaka, 1994; Grant, 1996), as well as with more recent studies highlighting the role of digital technologies in enabling distributed and collaborative knowledge processes.

AI systems operate as enabling infrastructures that enhance these processes by supporting pattern detection, facilitating interpretive interactions, and enabling continuous feedback loops. In this sense, AI does not replace human cognition but augments organizational knowledge processes by integrating algorithmic

reasoning and human judgment, contributing to the co-construction of meaning and the validation of knowledge within institutional contexts (Larsen, 2021; Leoni et al., 2024).

Surrounding these processes, the framework identifies a set of enabling conditions that shape the extent to which AI can function as a generative knowledge resource: 1) knowledge governance, 2) leadership, 3) organizational culture, and 4) capabilities. These conditions reflect existing insights from both knowledge management and public administration literature, which emphasize the importance of governance mechanisms, participatory practices, and cultural openness in supporting knowledge sharing and learning (Tangi et al., 2025). The transformation of knowledge processes leads to the development of organizational knowledge assets, including shared knowledge, validated insights, and institutional memory. This layer represents a critical conceptual advancement, as it shifts the focus from knowledge as a process to knowledge as an organizational resource that can be mobilized and leveraged in decision-making contexts.

Finally, the framework highlights key outcomes in terms of knowledge-based decision-making, adaptive capacity, and the development of more reflective and responsive public administrations. Importantly, the inclusion of feedback loops emphasizes the recursive nature of AI-enabled knowledge systems, in which outcomes continuously inform both the evolution of AI systems and the refinement of organizational knowledge processes.

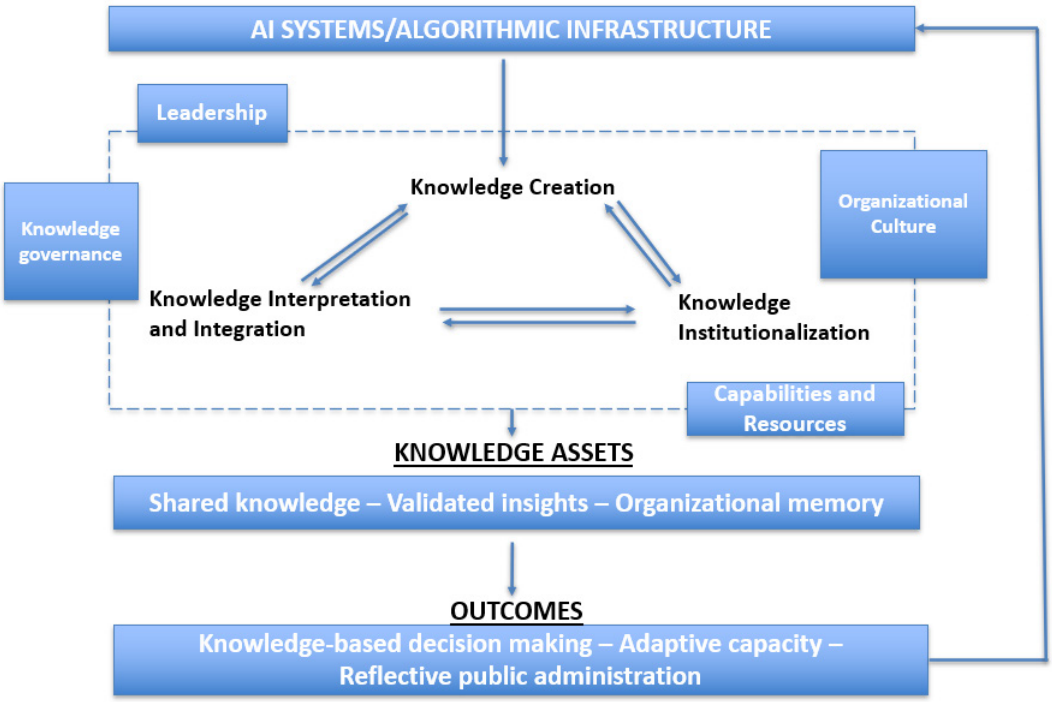


Figure 2: AI as a Knowledge Infrastructure in Public Administration

Building on the findings, the study also identifies a set of emerging research directions to further advance understanding of AI as a knowledge infrastructure in public administration.

First Research direction

Future research should explore how different configurations of AI systems influence specific knowledge processes, particularly the interplay between knowledge creation, integration, and institutionalization. Less is known, in fact, about how different technological architectures shape the quality, reliability, and usability of knowledge within organizations.

Second Research direction

There is a need to investigate the role of knowledge governance mechanisms in mediating the relationship between AI and organizational outcomes. In particular, future studies could examine how transparency, accountability, and participation affect the transformation of algorithmic outputs into shared and legitimate knowledge assets.

Third Research direction

Further attention should be devoted to the development of knowledge capabilities at both individual and organizational levels. This includes understanding how skills such as data literacy, interpretive capacity, and interdisciplinary collaboration contribute to the effective use of AI in knowledge-intensive contexts.

Fourth Research direction

Future research should explore the potential tensions between knowledge exploration and control, especially in highly regulated public environments. While AI enables experimentation and adaptive learning, institutional constraints may limit the development of open and reflexive knowledge processes.

Fifth Research direction

Longitudinal and comparative studies are needed to assess how AI-driven knowledge processes evolve over time and across different administrative contexts, contributing to a deeper understanding of how public organizations transition toward knowledge-based systems.

4.1 Theoretical implications

From a theoretical perspective, the present study advances knowledge management literature by integrating previously fragmented research streams, such as organizational learning, decision-making, and digital governance, into a unified framework that conceptualizes AI as a knowledge infrastructure embedded within organizational processes.

In doing so, the study not only consolidates existing theoretical perspectives but also proposes a structured research agenda that identifies key directions for future inquiry on AI-enabled knowledge processes in public administration.

Meta-synthesis has, in fact, pointed towards algorithms not just enhancing analytical capacity but also being involved in knowledge interpretation, integration, and transformation through interaction, feedback, and interpretive translation.

This insight calls for re-examining traditional assumptions about the rationality of public administration. The decision-making process is no longer constrained by human cognition or institutional routines alone but is increasingly enabled by hybrid knowledge processes emerging from the interaction between human judgment and algorithmic reasoning.

This entails a shift from a static notion of “technology adoption” toward a dynamic knowledge-based view of organizations, in which AI, data infrastructures, and human competencies jointly constitute evolving knowledge systems and knowledge assets.

The study therefore contributes to the literature by showing how AI enables the transformation of organizational knowledge processes—specifically knowledge creation, knowledge integration, and knowledge institutionalization, thereby reframing public administration as a system capable of continuously generating and reconfiguring its knowledge base.

4.2 Practical implications

From a practical perspective, the findings provide actionable insights for public managers and policymakers aiming to design and implement AI systems.

First, the findings emphasize the importance of governance architectures that balance surveillance and transparency, ensuring that AI systems function as shared knowledge assets rather than instruments of control. For instance, cross-functional data labs, steering boards, and participatory audit mechanisms can facilitate the transformation of algorithmic outputs into collectively validated organizational knowledge.

Second, capacity-building and leadership development emerge as critical levers. Leaders who promote dialogue, experimentation, and reflexivity foster knowledge-sharing cultures in which AI is used to support exploration and learning rather than compliance. In such contexts, errors and model revisions become opportunities for knowledge refinement and organizational learning.

Third, the study highlights the need for regulatory flexibility through frameworks that preserve ethical standards while enabling adaptive experimentation, such as regulatory sandboxes and iterative evaluation mechanisms (Ranchordas, 2021). These approaches support the continuous development and adjustment of AI-driven knowledge processes.

Finally, investments in interdisciplinarity and data literacy are essential to strengthen knowledge integration capabilities across different professional domains (Romeo et al., 2025). Enhancing collaboration between public officials, technical experts, and citizens enables the co-production and circulation of knowledge within public organizations.

Where these elements converge, AI evolves from a generative learning infrastructure to a generative knowledge infrastructure, capable of supporting continuous knowledge development and improving the adaptive capacity of public institutions.

5 Ethics Declaration

No ethical approval was required for this study. The research is based on a qualitative meta-synthesis of existing literature and does not involve human participants, personal data, or primary data collection.

6 AI Declaration

The authors declare that generative artificial intelligence (AI) tools, specifically ChatGPT, were used solely to improve the clarity and quality of the English language. As Italian is the authors' native language, AI support was limited to linguistic refinement. All intellectual content, analysis, and conclusions are entirely the responsibility of the authors.

References

Albous, M. R., & Alboloushi, B. (2025). AI-Driven Innovations in E-Government: How Is AI Reshaping the Public Sector?. In Harnessing AI, Blockchain, and Cloud Computing for Enhanced e-Government Services (pp. 93-118). IGI Global Scientific Publishing.

Babšek, M., Ravšelj, D., Umek, L., & Aristovnik, A. (2025). Artificial intelligence adoption in public administration: An overview of top-cited articles and practical applications. AI, Vol. 6 No. 3, p. 44.

Bolton, M., Raven, R., & Mintrom, M. (2021). Can AI transform public decision-making for sustainable development? An exploration of critical earth system governance questions. *Earth System Governance*, Vol.9 No. 100116.

Camilleri, M. A. (2024). Artificial intelligence governance: Ethical considerations and implications for social responsibility. *Expert systems*, Vol. 41 No. 7, p. e13406.

Cetina, K. K. (2007). Culture in global knowledge societies: Knowledge cultures and epistemic cultures. *Interdisciplinary science reviews*, Vol. *32* No.4, pp. 361-375.

Giraldi, L., Rossi, L., & Rudawska, E. (2024). Evaluating public sector employee perceptions towards artificial intelligence and generative artificial intelligence integration. *Journal of Information Science*, Vol. 01655515241293775.

Grant, R. M. (1996). Toward a knowledge-based theory of the firm. *Strategic management journal*, Vol. *17* No. S2, pp. 109-122.

Janssen, M., Konopnicki, D., Snowdon, J. L., & Ojo, A. (2017). Driving public sector innovation using big and open linked data (BOLD). *Information systems frontiers*, Vol. *19* No. 2, pp. 189-195.

Laihonen, H., & Kokko, P. (2023). Knowledge management and hybridity of institutional logics in public sector. *Knowledge management research & practice*, Vol. *21* No. 1, pp. 14-28.

Larsen, B. C. (2021). A framework for understanding AI-induced field change: How AI technologies are legitimized and institutionalized. In Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society (pp. 683-694).

Leoni, L., Gueli, G., Ardolino, M., Panizzon, M., & Gupta, S. (2024). AI-empowered KM processes for decision-making: empirical evidence from worldwide organisations. *Journal of Knowledge Management*, Vol. 28 No. 11, pp. 320-347.

Mainardi, I. (2025). Change management: artificial intelligence (AI) at the service of public administrations. *AI & SOCIETY*, Vol. 40 No. 5, pp. 3953-3981.

Nonaka, I. (1994). A dynamic theory of organizational knowledge creation. Organization science, Vol. 5 No. 1, pp. 14-37.

Noah Jr, P. D. (2017). A Systematic Approach To The Qualitative Meta-Synthesis. *Issues in Information Systems*, Vol. 18 No. 2.

Olsen, H. P., Hildebrandt, T. T., Wiesener, C., Larsen, M. S., & Flügge, A. W. A. (2024). The right to transparency in public governance: Freedom of information and the use of artificial intelligence by public agencies. *Digital Government: Research and Practice*, Vol. *5* No. 1, pp. 1-15.

Oryshchuk, V. A. S. Y. L., & Minenko, Y. E. H. O. R. (2024). Using AI to optimize decision-making processes in public administration. *Armenian Journal of Public Administration*, Vol. 1 No. 24, pp. 38-43.

Patriotta, G. (2004). On studying organizational knowledge. *Knowledge Management Research & Practice*, Vol. *2* No. 1, pp. 3-12.

Ranchordas, S. (2021). Experimental regulations for AI: sandboxes for morals and mores. *Morals & Machines*, Vol. *1* No. 1, pp. 86-100.

Romeo, E., Giampaola, V., & Capolupo, N. (2025). *Oltre i dati. Come la cultura e le capacità organizzative abilitano la business model innovation* (pp. 1-161). Giappichelli.

Tangi, L., Müller, A. P. R., & Janssen, M. (2025). AI-augmented government transformation: Organisational transformation and the sociotechnical implications of artificial intelligence in public administrations. *Government Information Quarterly*, Vol. 42 no. 10205, p. 5.

Walsh, D., & Downe, S. (2005). Meta-synthesis method for qualitative research: a literature review. *Journal of advanced nursing*, Vol. 50 No. 2, pp. 204-211.

Weinberger, M. (2022). What is metaverse? —A definition based on qualitative meta-synthesis. *Future Internet*, Vol. 14 No. 11, pp. 310.

The Social Cost of Digital Banking: Branch Desertification and Inequalities in Access to Financial Services

Giuseppe Bruno, Carmela Piccolo, Eduardo Pipicelli
University of Naples Federico II, Naples, Italy

Abstract

In recent years, the banking sector has undergone a profound transformation driven by the widespread diffusion of digital technologies, which has progressively reshaped service delivery models and reduced the role of physical branches. This shift has been accompanied by a significant contraction of branch networks across Europe, raising concerns about its implications for access to financial services, particularly in less advantaged territorial contexts. Against this background, this study investigates the social costs associated with the digitalization of banking services by analyzing how branch network contraction affects spatial accessibility across different territorial contexts. The empirical analysis focuses on a regional case study in Italy, Emilia-Romagna, characterized by a high degree of heterogeneity in socio-economic and digital conditions. Accessibility is assessed through two complementary dimensions: the *availability* of branches within given distances and the *proximity* to the nearest facility. The analysis is conducted for three reference years (2010, 2017, and 2024), enabling a detailed comparison of the evolution of accessibility patterns across urban, suburban, and rural areas. The results show that, although branch closures have occurred throughout the region, their impact has been significantly more pronounced in suburban and rural areas, where both availability and proximity indicators reveal a marked deterioration of accessibility conditions. Building on this evidence, the study further explores the relationship between physical accessibility and digital development, showing that the areas most affected by the reduction in physical infrastructure are also characterized by weaker digital conditions. In these contexts, digital channels do not adequately compensate for the loss of physical access, leading to a cumulative effect of disadvantage. Overall, the findings suggest that the ongoing digital transformation of the banking sector entails non-negligible social costs, as it disproportionately affects more vulnerable populations and contributes to widening territorial disparities in access to financial services. The study highlights the need to account for spatial and digital inequalities when designing policies aimed at promoting financial inclusion in the context of digital transition.

Keywords: Banking desertification; Digitalization; Accessibility; Financial Inclusion, Social Sustainability.

Paper type: Academic Research Paper

1 Introduction and Background

Over the last few decades, the banking sector has undergone a profound transformation driven by the combined effects of financial crises and rapid technological innovation (Felipe et al., 2025; Kitsios et al., 2021). The diffusion of online banking platforms has progressively reshaped the organization of financial activities, encouraging institutions to move toward increasing digital business models (Thakor, 2020). Within this transformation, a growing share of banking services is performed through remote interfaces, reducing the centrality traditionally played by physical branches. For this reason, many banking institutions have progressively rationalized their branch networks in order to reduce operational costs and adapt their distribution strategies to the new environment (Baldassarre et al., 2024).

As a consequence, the number of bank branches has declined significantly in many European countries, contributing to the emergence of bank branch desertification (Langford et al., 2024). While this process is often interpreted as the result of technological progress and efficiency-oriented strategies, it has also raised concerns regarding its implications for financial inclusion, particularly in rural and remote areas (Baldassarre et al., 2025; Demirgüç-Kunt et al., 2021).

The extent to which physical banking services can be substituted by digital channels depends on factors such as connectivity and the diffusion of digital capabilities, which vary significantly across territories and socioeconomic groups (Calzada & Pablo, 2026). As a result, in contexts where these conditions are less developed, the contraction of branch networks may reduce effective access to financial services, generating non-negligible social costs and reinforcing existing territorial disparities (Bruno et al., 2023).

Against this background, an increasing body of literature on banking digitalization has examined the implications of branch network contraction for access to the service, with particular attention to financial inclusion and territorial inequalities (Leyshon et al., 2008; Hegerty, 2016). In this context, several contributions have analyzed spatial accessibility using quantitative and spatial approaches, showing how changes in network dimensions translate into variations in physical access to banking services (Galardo et al., 2021; Higgs et al.,

2022). These analyses consistently point to the emergence of uneven access conditions, as branch closures tend to disproportionately affect already disadvantaged areas, contributing to the persistence and, in some cases, the intensification of territorial disparities (Hegerty, 2020; Langford et al., 2021; Pardue & Shelton, 2025).

While existing contributions document the spatial implications of branch network contraction, limited attention has been devoted to empirically assessing how these effects impact more fragile areas, with lower structural and digital conditions. This aspect is crucial, as in such contexts the reduction of physical access may not be offset by digital channels, compounding existing disadvantages and ultimately exacerbating territorial divides (de Ibarreta et al., 2025).

Building on this perspective, the present study aims to investigate the social costs of banking digitalization by addressing the following research question (RQ): *How does bank branch desertification, driven by digitalization, exacerbate territorial inequalities in access to financial services?*

To address this research question, the analysis focuses on a regional case study in Italy, namely Emilia-Romagna, a context characterized by significant branch desertification and a heterogeneity in economic and digital conditions across urban, suburban, and rural areas. Within this setting, the study evaluates spatial accessibility to bank branches (measured in terms of availability and proximity) across different territorial contexts, tracking its evolution over approximately fifteen years, with a focus on the years 2010, 2017, and 2024.

The results show that, although branch closures have been distributed across the entire regional territory, their impact has been more pronounced in rural (and suburban) areas, where they have led to a more marked deterioration of accessibility conditions. These effects are particularly critical, as lower levels of digital development in these areas are often insufficient to compensate for reductions in physical infrastructure, thereby exacerbating existing disparities.

The remainder of the paper is organized as follows. Section 2 introduces the Italian case study, with a focus on bank branch desertification in the Emilia-Romagna region. Section 3 presents the methodology adopted for the spatial accessibility analysis and discusses the corresponding results. Section 4 provides additional insights into the relationship between physical and digital accessibility, highlighting its implications for territorial inequalities. Finally, Section 5 concludes the paper by discussing its main limitations and outlining directions for future research.

2 Bank Branch Desertification in Italy and the Emilia-Romagna Case Study

Over the last decade, the Italian banking system has undergone a progressive reorganization of its territorial structure, reflected in a widespread reduction in the number of physical branches. This process has affected all regions, although with different intensities, and represents one of the most visible outcomes of the ongoing transformation of banking distribution models.

Table 1 reports the evolution of the number of bank branches across Italian regions between 2010 and 2024, together with the corresponding percentage variations (Banca di Italia, 2025). The data show a generalized contraction of branch networks. At the national level, the number of branches declined from 33,663 units in 2010 to 19,655 in 2024, corresponding to a reduction of approximately 42%. In general, the closures appear consistent across the country, although their intensity slightly varies by region. In several cases, the reduction exceeds 45%, as observed in regions such as Marche, Molise, Umbria, Veneto, and Emilia Romagna.

Table 1: Number of branches per region

Region	Population	Number of branches		
		2010	2017	2024
Abruzzo	1,272,627	701	587 (−16%)	400 (−43%)
Basilicata	537,577	250	226 (−10%)	163 (−35%)
Calabria	1,846,610	519	422 (−19%)	315 (−39%)
Campania	5,609,536	1,637	1,368 (−16%)	1,008(−38%)
Emilia-Romagna	4,425,366	4,149	3,357 (−19%)	2,073 (−50%)
Friuli-Venezia Giulia	1,194,248	955	776 (−19%)	560 (−41%)
Lazio	5,720,536	2,768	2,294 (−17%)	1,607 (−42%)
Liguria	1,507,636	962	767 (−20%)	560(−42%)
Lombardia	9,976,509	6,611	5,431 (−18%)	3,769 (−43%)
Marche	1,484,298	1,206	904 (−25%)	613 (−49%)
Molise	290,636	146	119 (−18%)	76 (−48%)
Piemonte	4,251,351	2,696	2,252 (−16%)	1,574 (−42%)
Puglia	3,907,683	1,410	1,203 (−15%)	923 (−35%)
Sardegna	1,578,146	667	568 (−15%)	454 (−32%)
Sicilia	4,814,016	1,759	1,416 (−19%)	1,030 (−41%)

Region	Population	Number of branches		
		2010	2017	2024
Toscana	3,661,981	2,548	2,093 (−18%)	1,472 (−42%)
Trentino-Alto Adige	1,077,143	979	789 (−19%)	658 (−33%)
Umbria	856,407	581	458 (−21%)	317 (−45%)
Valle d'Aosta	12,313	98	93 (−5%)	66 (−33%)
Veneto	4,849,553	3,625	2,801 (−23%)	2,007 (−45%)
Italy	58,997,201	33,663	27,374 (−19%)	19,655 (−42%)

These results indicate that the contraction of the bank branch network is a systemic process affecting the entire country; however, in more densely populated and economically developed areas, the reduction may have limited consequences. By contrast, in more peripheral contexts, it can contribute to the emergence of bank branch desertification, potentially leading to financial exclusion when the loss of physical infrastructure is not compensated by alternative channels.

Against this background, we focus our analysis on the Emilia-Romagna region, which experienced the highest percentage reduction in the country (−50%) and is characterized by heterogeneity in economic and digital conditions across different areas.

Beyond population distribution, the region's socio-economic structure also reflects territorial differences, as reported in Table 2 and illustrated in Figures 1c and 1d, based on data from Regione Emilia-Romagna (2025). In particular, per capita income tends to decrease from urban to suburban and rural areas, with a gap of about 4,000 euros between the former and the latter. This pattern is further reinforced by the distribution shown in Figure 1d, which reports the share of taxpayers with total income below €10,000. Higher concentrations of low-income individuals are observed in peripheral and rural municipalities, while urban areas display systematically lower shares. These outcomes highlight a socio-economic divide across the region, with more fragile conditions concentrated outside the main urban centres.

With respect to the distribution of bank branches, the reduction follows a broadly similar pattern across areas. As noted, the region experienced a substantial contraction of approximately 50% between 2010 and 2024, which appears relatively homogeneous across territorial contexts, as also suggested by the spatial distribution of branches in Figure 1b. However, while urban areas still benefit from a relatively dense network (791), in rural contexts the same reduction translates into much more limited branch availability for potential users (449).

Taken together, these elements point to the need for a more careful assessment of accessibility conditions across the region. In fact, while the contraction of the networks is broadly distributed, its effects on financial exclusion are likely to be more critical in suburban and rural areas that are economically weaker.

Table 2: Key characteristics of Emilia-Romagna by area

Area	Municipalities		Census tracts		Population		Per capita income (Euro)	No. of branches		
	#	%	#	%	#	%		2010	2017	2024
Urban	11	3%	17,001	29,231.7	1,703,827	39%	29,231.7	1746	1370 (-22%)	791 (-55%)
Suburban	112	34%	12,626	26,907.3	1,804,646	41%	26,907.3	1620	1310(-19%)	833 (-49%)
Rural	207	63%	11,864	25,268.9	916,893	21%	25,268.9	783	677 (-14%)	449 (-43%)
Total	330	100%	41,491	27,462.8	4,425,366	100%	27,462.8	4,149	3,357 (-19%)	2,073 (-50%)

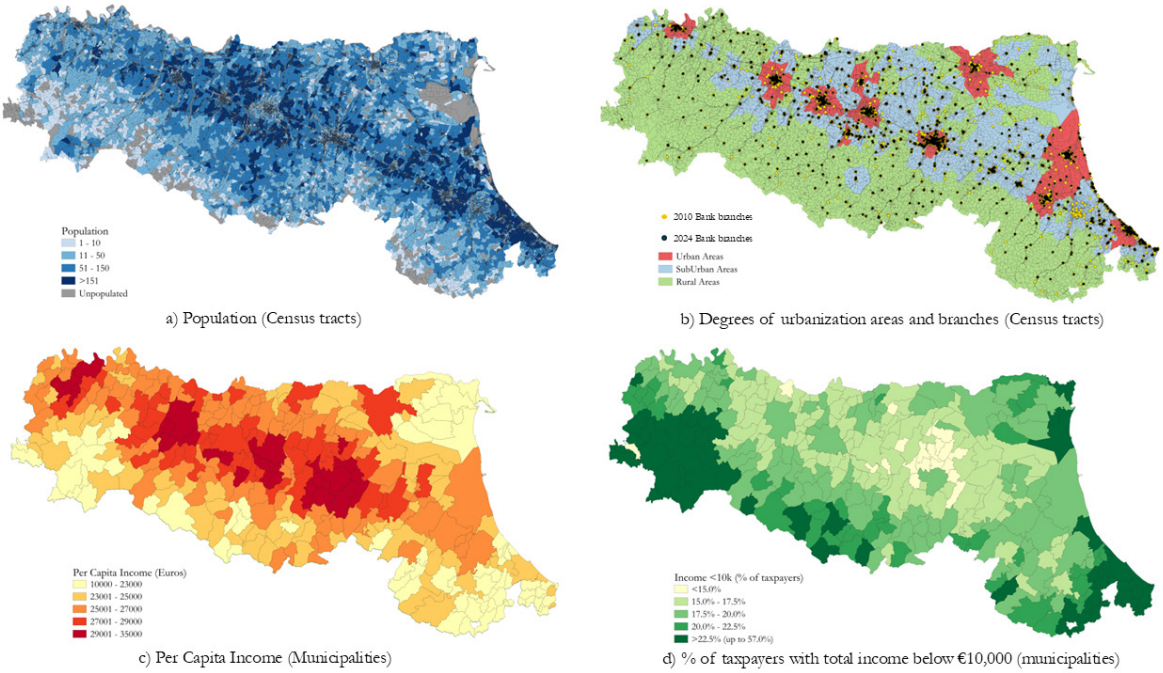


Figure 1: Territorial distribution of key characteristics in Emilia-Romagna

3 Spatial Accessibility Analysis

This section presents the empirical analysis of spatial accessibility to banking services. It begins with a description of the methodological framework (Section 3.1) and then discusses the main results, with a focus on how accessibility conditions evolve across different territorial contexts (Section 3.2).

3.1 Methodology

Spatial accessibility is commonly defined as the ease with which individuals can reach a given service, and it has been widely examined in the literature through different methodological approaches (Neutens, 2015). Among these, distance-based measures are among the most widely adopted, as they provide a consistent way to capture spatial interaction between demand and supply.

In this study, accessibility to banking services is assessed considering two metrics: *availability* of alternative branches within given distances and *proximity* to the closest facility. The former reflects the range of opportunities available to users within given distance thresholds, while the latter captures the minimum effort (in terms of distance) required to access financial services.

From an interpretative perspective, considering both availability and proximity is essential to properly assess access conditions to banking services. Availability captures the number of alternatives accessible within a given distance, thus reflecting the variety of options available to users. This aspect is particularly relevant for services such as banking, where customers may differ in their preferences across institutions in terms of costs, services, or existing relationships. However, access is not only a matter of choice, but also of reachability. In this sense, proximity complements availability by capturing the distance that users must cover to reach the closest branch, providing a direct indication of physical access.

The analysis is conducted at the level of census tracts, which are used to discretize the spatial distribution of demand, as each tract is represented by its centroid, approximating the location of the resident population. Let I denote the set of demand locations, corresponding to census tract centroids, and let J^tbe the set of active bank branches at time $t \in \{2010,2017,2024\}$. Each location $i \in I$ is associated with a demand u_i, representing the number of potential users and defined as individuals aged between 20 and 75 years. Bank branches are georeferenced, and distances d_{ij}between demand nodes and facilities are computed as shortest paths along the road network using the open-source software QGIS.

The first dimension of accessibility is availability, which captures the number of alternative branches that can be reached within a given distance. For each location $i \in I$and a distance threshold $\bar{d}$, availability is defined as $A_i^t(\bar{d}) = |\{j \in J^t : d_{ij} \leq \bar{d}\}|$.

From this measure, two indicators are derived as follows.

- The fraction of users with at least a given number of opportunities $k$within a distance threshold $\bar{d}$:

$$G_k^t(\bar{d}) = \frac{\sum_{i \in I: A_i^t(\bar{d}) \geq k} u_i}{\sum_{i \in I} u_i}. \qquad (1)$$

- The average number of branches accessible to potential users within the given distance threshold $\bar{d}$:

$$A_{\text{avg}}^t(\bar{d}) = \frac{\sum_{i \in I} u_i \, A_i^t(\bar{d})}{\sum_{i \in I} u_i}. \qquad (2)$$

The second dimension of accessibility is proximity, defined as the minimum distance between each demand location and the nearest active branch. Formally, for each $i \in I$, proximity is measured as $d_i^t = \min_{j \in J^t} \{d_{ij}\}$.

From this, two indicators are derived as follows.

- The fraction of users that can reach the closest branch within a predefined distance $\bar{d}$:

$$F^t(\bar{d}) = \frac{\sum_{i \in I: d_i^t \leq \bar{d}} u_i}{\sum_{i \in I} u_i}. \qquad (3)$$

- The average distance that a potential user must travel to reach the closest branch:

$$d_{\text{avg}}^t = \frac{\sum_{i \in I} u_i \, d_i^t}{\sum_{i \in I} u_i}. \qquad (4)$$

On the basis of these measures, the analysis is then carried out on the case study, allowing us to assess how accessibility conditions evolve across different territorial contexts over time, as presented in the next section.

3.2 Results

The results are presented with respect to the indicators introduced in Section 3.1, considering the three reference years (2010, 2017, and 2024) and distinguishing across the three territorial contexts (urban, suburban, and rural areas).

We begin by discussing availability, reported in Table 3. For the sake of clarity, we specify that results are presented with reference to a distance threshold of 2.0 km, which represents a reasonable compromise between very short distances, often too restrictive, and larger thresholds that tend to smooth out territorial differences.

The table reports the fraction of users who can access at least k branches within this distance, G_k^t, together with the corresponding average number of opportunities, A_{avg}^t.

Over time, the contraction of branch networks significantly reduces the number of accessible opportunities, especially in suburban and rural contexts. For instance, the share of users with access to at least three branches within 2.0 km declines by 18% in suburban areas and by 39% in rural areas, compared to a reduction of 7% in urban areas. A similar pattern emerges when considering higher levels of availability. When considering four accessible branches, the reduction reaches 20% in suburban areas and 58% in rural areas, while remaining at 8% in urban contexts. This dynamic is also reflected in the average number of available branches, which decreases from 43 to 19 in urban areas, from 10 to 5 in suburban areas, and from 3 to 2 in rural areas.

As a result, by 2024, availability conditions markedly differ across territorial contexts. In urban areas, a large share of the population still has access to multiple branches within 2 km, preserving a relatively high degree of choice. By contrast, in suburban and especially rural areas, the reduction in opportunities leads to much lower levels of availability, with several users having access to only a few branches.

Table 3: Availability results: *fraction of users* (G_k^t, in %) by k and average distance A_{avg}^t ($\bar{d} = 2.0\ km$)

			No. of branches (k)										
$\bar{d}$	**Area**	**Year (t)**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10 or more**	A_{avg}^t
2.0 km	Urban	2010 (a)	92%	87%	83%	81%	79%	78%	76%	75%	73%	72%	43
		2017 (b)	90%	86%	82%	79%	77%	75%	73%	71%	69%	67%	33
		2024 (c)	87%	81%	77%	74%	71%	67%	64%	61%	59%	57%	19
		Δ [%] ($\frac{c-a}{a} * 100$)	-5%	-6%	-7%	-8%	-10%	-14%	-16%	-18%	-20%	-21%	24
	Suburban	2010 (a)	84%	78%	73%	67%	62%	57%	52%	49%	40%	34%	10
		2017 (b)	82%	76%	70%	65%	59%	53%	47%	39%	32%	29%	8
		2024 (c)	75%	67%	60%	54%	47%	36%	28%	25%	21%	18%	5
		Δ [%] ($\frac{c-a}{a} * 100$)	-11%	-14%	-18%	-20%	-24%	-36%	-46%	-50%	-48%	-47%	5

			No. of branches (k)										
$\bar{d}$	Area	Year (t)	1	2	3	4	5	6	7	8	9	10 or more	A^t_{avg}
	Rural	2010 (a)	68%	56%	39%	30%	21%	14%	8%	5%	4%	2%	3
		2017 (b)	67%	51%	35%	25%	16%	9%	6%	4%	2%	0%	2
		2024 (c)	60%	41%	24%	13%	6%	4%	2%	1%	0%	0%	2
		Δ [%] ($\frac{c-a}{a} * 100$)	-11%	-26%	-39%	-58%	-72%	-71%	-69%	-84%	-100%	-100%	1

Table 4: Proximity results: *fraction of users* (F^t, in %) by $\bar{d}$ and average distance d^t_{avg}

		Distance [km]					
Area	Year (t)	1.0	2.0	3.0	4.0	5.0	d^t_{avg}[km]
Urban	2010 (a)	82%	92%	96%	98%	99%	0.7
	2017 (b)	79%	90%	95%	98%	99%	0.8
	2024 (c)	73%	87%	92%	96%	98%	1.0
	Δ [%] ($\frac{c-a}{a} * 100$)	-10%	-5%	-4%	-3%	-1%	+0.3
Suburban	2010	65%	84%	91%	96%	98%	1.1
	2017	63%	82%	90%	95%	98%	1.2
	2024	54%	75%	83%	88%	92%	1.6
	Δ [%] ($\frac{c-a}{a} * 100$)	-17%	-11%	-9%	-8%	-6%	+0.5
Rural	2010	53%	68%	77%	85%	90%	1.9
	2017	51%	67%	76%	84%	90%	2.0
	2024	45%	60%	69%	79%	86%	2.3
	Δ [%] ($\frac{c-a}{a} * 100$)	-15%	-11%	-10%	-7%	-5%	+0.4

These results highlight a substantial reduction in the range of accessible opportunities, particularly in suburban and rural areas. However, to provide a more complete picture of accessibility, we also consider the distance that potential customers must travel to reach the nearest branch, as captured by the proximity indicators (Table 4).

The latter reports the fraction of users that can reach the closest branch within increasing distance thresholds F^t, together with the corresponding average distance d^t_{avg}.

A first relevant aspect is that, even before the contraction of branch networks, accessibility conditions were already markedly different across territorial contexts. In 2010, for instance, 82% of the urban population could reach a branch within 1 km, compared to 65% in suburban areas and 53% in rural ones, with a clear initial gap in proximity conditions.

The reduction of branches over time has further amplified these differences. In fact, although closures have been broadly distributed across the region (see Table 2), their relative impact appears more pronounced in suburban and rural areas, which were already characterized by less favorable conditions. For example, within 1 km, the share of covered population in urban areas decreased by 10%, reaching 73% in 2024. By contrast, the reduction amounts to 17% in suburban areas and 15% in rural areas, leading to significantly lower coverage levels of 54% and 45%, respectively.

Similar patterns emerge across higher distance thresholds and in the average distance, where comparable increases translate into more critical accessibility conditions in suburban and rural areas.

4 Additional Insights: Digital Divide and Territorial Inequalities

The evidence discussed in Section 3 highlights a more pronounced deterioration of physical accessibility in suburban and rural areas. Clearly, given that the contraction of branch networks is largely driven by the digitalization of banking services, a key question is whether areas more exposed to this process are characterized by adequate levels of digital development.

To address this aspect, we consider two complementary indicators, retrieved from Regione Emilia-Romagna (2026): (i) the share of households covered by FTTH (Fiber to the Home), which captures the level of development of digital infrastructure; and (ii) the share of citizens with a digital identity, which reflects the

degree of digital familiarity and readiness to access online services. Together, these measures provide a synthetic view of both the infrastructural and skills dimensions of digital access.

The spatial distribution of these indicators, reported in Figure 2, reveals a clear territorial pattern. As shown in Figure 2a, FTTH coverage is higher in urban municipalities, while suburban and especially rural areas exhibit lower levels of infrastructure deployment, with several peripheral zones still characterized by limited connectivity. Indeed, on average, FTTH coverage reaches around 56% in urban areas, compared to approximately 20% in suburban areas and 22% in rural ones, pointing to a substantial infrastructural gap.

A similar result emerges when considering digital identity adoption, reported in Figure 2b. While the differences across areas are less pronounced than for infrastructure, urban contexts still display higher levels of digital engagement. The average share of citizens with a digital identity is approximately 36% in urban areas, compared to 34% in suburban areas and 29% in rural ones. Again, the lower values observed in rural areas suggest a weaker capacity to effectively rely on digital services.

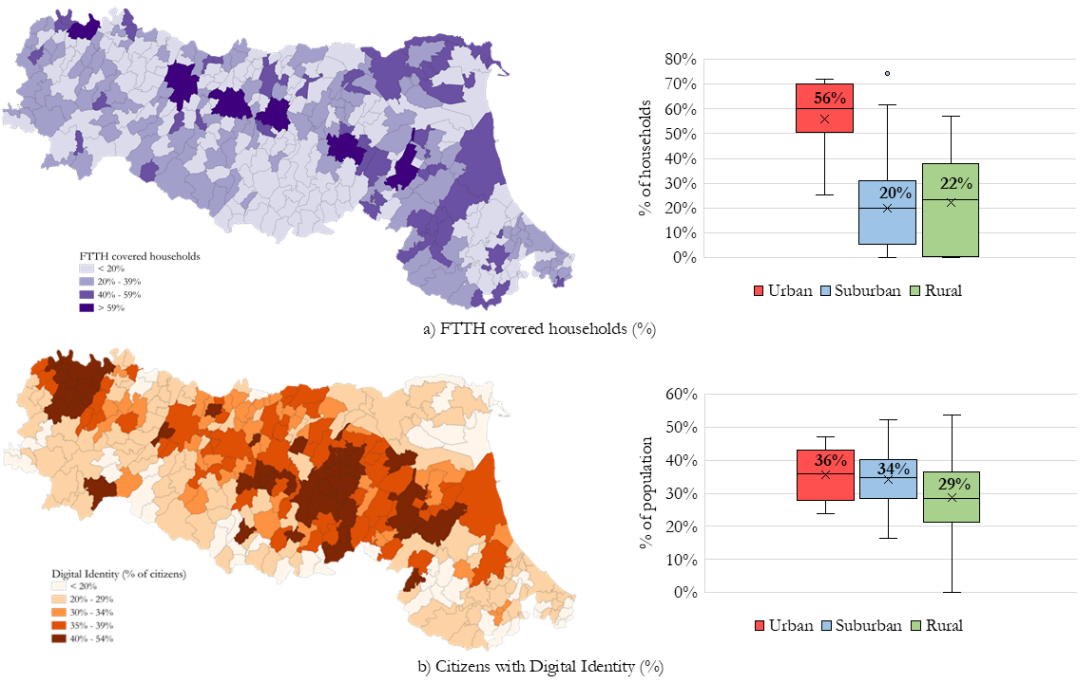


Figure 2: Digital infrastructure and digital identity indicators across municipalities

Taken together, these results indicate that the territories most affected by the reduction in physical accessibility are also those characterized by less favorable digital conditions. This represents a particularly relevant insight, as in urban areas the (contained) reduction of physical infrastructure is partially offset by relatively high levels of digital development, whereas in suburban and rural contexts this does not occur. Rather, as the quality of physical access to banking services deteriorates, the ability to compensate for this loss through alternative digital channels remains limited.

As a result, what is essentially a process of digital transformation in the banking sector entails a non-negligible social cost, mainly stemming from its disproportionate impact on more vulnerable populations, whose access to services worsens as it increasingly depends on digital channels to which they are comparatively less equipped to adapt.

5 Conclusions

The ongoing digital transformation of the banking sector is reshaping the way financial services are delivered, progressively reducing the role of physical infrastructure. While this process is often associated with efficiency gains and improved service provision, its implications are not evenly distributed across territories.

In this study, we investigated the social costs associated with this transition by analyzing how the contraction of branch networks affects accessibility to banking services across different territorial contexts. Focusing on the Emilia-Romagna region, the empirical analysis shows that, although branch closures have occurred throughout the entire territory, their impact has been significantly more pronounced in suburban and rural areas. These

contexts, already characterized by less favorable accessibility conditions, experienced a more substantial deterioration both in terms of availability of services and proximity to the nearest branch.

Building on this evidence, the analysis further highlights that such patterns are closely intertwined with the uneven distribution of digital development across the region. In particular, the areas most affected by the reduction in physical accessibility are also those with comparatively weaker digital infrastructure and capabilities. As a result, the transition toward digital service provision does not serve as a compensatory mechanism but rather amplifies existing territorial disparities.

From this perspective, the main contribution of the paper lies in providing empirical evidence of a cumulative effect of disadvantage for more vulnerable populations. In this sense, the digitalisation of banking services entails a non-negligible social cost, as it strongly reinforces inequalities in access to essential financial services.

These findings have relevant policy implications for public stakeholders committed to reducing territorial disparities. In particular, they suggest that the ongoing digital transformation of the banking sector should be accompanied by targeted interventions to support more fragile areas, including investments in digital infrastructure, initiatives to enhance digital skills, and alternative service delivery solutions capable of ensuring minimum access conditions in contexts affected by branch desertification.

Clearly, the study presents some limitations that open to future research. First, the analysis focuses on a single regional case study, and further investigations could extend the approach to other contexts to assess the generalizability of the results. Second, while accessibility is measured in spatial terms, future research could integrate behavioral dimensions, such as actual usage of digital banking services, to better capture effective access. Moreover, exploring the role of alternative financial channels could provide additional insights into how to mitigate the emerging gaps identified in this study.

6 Ethics declaration

The authors declare that they have complied with all ethical requirements. No ethical approval was required. The authors declare no conflict of interest.

7 AI declaration

The authors used AI-based tools to support proofreading and language editing. All content, analyses, and interpretations presented in the paper are the original work of the authors, who take full responsibility for the final version of the manuscript.

References

Baldassarre, S., Bruno, G., Piccolo, C., & Ruiz-Hernández, D. (2024). Multi-channel distribution in the banking sector and the branch network restructuring. Expert Systems with Applications, 238, 122294.

Baldassarre, S., Cavola, M., & Pipicelli, E. (2025). Bank branch desertification and the design of a postal banking network. Socio-Economic Planning Sciences, 100, 102216.

Banca d'Italia (2025). *Banks and financial institutions in Italy*. URL: https://www.bancaditalia.it/statistiche, (Accessed on 25 April 2026).

Bruno, G., Diglio, A., Piccolo, C., & Pipicelli, E. (2023). A reduced Composite Indicator for Digital Divide measurement at the regional level: An application to the Digital Economy and Society Index (DESI). Technological Forecasting and Social Change, 190, 122461.

Calzada, J., & Pablo, A. (2025). Broadband deployment and access to online banking: Does improved technology facilitate adoption?. Telecommunications Policy, 103113.

de Ibarreta, C. M., Ruíz-Rúa, A., Gijón, C., & Fernández-Bonilla, F. (2025). Banking digitalization in Spain: How branch closures and digital barriers reshape financial inclusion. Digital Business, 5(1), 100119.

Demirgüç-Kunt, A., Klapper, L., Singer, D., & Ansar, S. (2021). Financial inclusion, digital payments, and resilience in the age of COVID-19. World Bank Report.

Felipe, T., de Oliveira, R. T., Toth-Peter, A., Mathews, S., & Dulleck, U. (2025). Digital transformation in commercial banks: Unraveling the flow of Industry 4.0. Digital Business, 5(2), 100129.

Galardo, M., Garrì, I., Mistrulli, P. E., & Revelli, D. (2021). The geography of banking: Evidence from branch closings. Economic Notes, 50(1), e12177.

Hegerty, S. W. (2016). Commercial bank locations and "banking deserts": A statistical analysis of Milwaukee and Buffalo. The Annals of Regional Science, 56(1), 253-271.

Hegerty, S. W. (2020). "Banking deserts," bank branch losses, and neighborhood socioeconomic characteristics in the City of Chicago: A spatial and statistical analysis. The Professional Geographer, 72(2), 194-205.

Higgs, G., Price, A., & Langford, M. (2022). Investigating the impact of bank branch closures on access to financial services in the early stages of the COVID-19 pandemic. Journal of Rural Studies, 95, 1–14.

ISTAT (2026). *Basi territoriali e variabili censuarie*. URL: https://www.istat.it/notizia/basi-territoriali-e-variabili-censuarie/, (Accessed on 25 April 2026).

Kitsios, F., Giatsidis, I., & Kamariotou, M. (2021). Digital transformation and strategy in the banking sector: Evaluating the acceptance rate of e-services. Journal of Open Innovation: technology, market, and complexity, 7(3), 204.

Langford, M., Higgs, G., & Jones, S. (2021). Understanding spatial variations in accessibility to banks using variable floating catchment area techniques. Applied Spatial Analysis and Policy, 14(3), 449–472.

Langford, W. S., Thomas, H. W., & Feldman, M. P. (2024). Banking for the other half: The factors that explain banking desert formation. Economic Development Quarterly, 38(2), 71-81.

Leyshon, A., French, S., & Signoretta, P. (2008). Financial exclusion and the geography of bank and building society branch closure in Britain. Transactions of the Institute of British Geographers, 33(4), 447-465.

Neutens, T. (2015). Accessibility, equity and health care: review and research directions for transport geographers. *Journal of Transport Geography*, *43*, 14-27.

Pardue, R., & Shelton, T. (2025). Banks on a plane: The possibilities and limits of spatial accessibility measures for understanding geographies of financial exclusion. Environment and Planning A: Economy and Space, 57(2-3), 325-345.

Regione Emilia-Romagna (2025). *La potenziale fragilità demografica, sociale ed economica nei comuni dell'Emilia-Romagna nel 2023*. URL: https://statistica.regione.emilia-romagna.it/studi-analisi/2025/mappe-potenziale-fragilita-emilia-romagna-2023, (Accessed on 25 April 2026).

Regione Emilia-Romagna (2026). *DESIER – Digital Economy and Society Index Emilia-Romagna*. URL: https://digitale.regione.emilia-romagna.it/progetti/desier, (Accessed on 25 Apr 2026).

Thakor, A. V. (2020). Fintech and banking: What do we know?. Journal of financial intermediation, 41, 100833.

Assessing VSME Alignment in Voluntary Sustainability Reporting by SMEs: Evidence from the Italian Chemical Sector

Manuel Cavola[1], Patrice De Micco[2], Niccolò Parissi[2]
[1] Università Telematica Pegaso, Napoli, Italy
[2] Università degli Studi di Siena, Siena, Italy

Abstract

Small and medium-sized enterprises (SMEs) play a central role in the European transition towards a more sustainable economy, yet their sustainability reporting practices remain less developed and less standardised than those of large firms. This issue is particularly relevant in the context of the Corporate Sustainability Reporting Directive (CSRD), the European Sustainability Reporting Standards (ESRS), and the increasing sustainability-related information demands that SMEs face from large companies, financial institutions, and other stakeholders. In response to proportionality concerns, the EFRAG Voluntary Sustainability Reporting Standard for non-listed SMEs (VSME) provides a simplified framework intended to support voluntary sustainability disclosure by SMEs. However, empirical evidence on the extent to which SME sustainability reports are aligned with the VSME remains limited.

This paper assesses the alignment of voluntary sustainability reporting by SMEs operating in the Italian chemical sector with VSME disclosure requirements. The empirical analysis focuses on firms classified under the C20 division of the NACE industrial classification. Starting from a population of Italian chemical firms identified through the AIDA database, the study selects SMEs that had published a sustainability report or an equivalent sustainability-related disclosure document. A stratified sampling strategy was then adopted, and the final analytical dataset was obtained after a quality-control procedure on document extraction.

The analysis applies an AI-assisted content analysis protocol based on a VSME coding grid composed of 20 disclosure items, covering the Basic and Comprehensive Modules and four thematic areas: General information, Environment, Social, and Governance. Each firm-item pair was assessed through a three-level scoring scale distinguishing between no alignment, partial alignment, and full substantive alignment with the VSME requirements.

The results show a moderate and uneven degree of alignment with the VSME. Full and partial alignment are both relatively frequent, while non-coverage remains substantial but not dominant. At firm level, the findings indicate relevant heterogeneity across companies. General information records the highest alignment, followed by Social and Environment, whereas Governance is the weakest area. At item level, firms perform better on business-model descriptions, sustainability initiatives, and human-rights policies, while weaker alignment emerges for biodiversity, pollution, climate-transition targets, corruption-related sanctions, and sensitive-sector revenues. Overall, the findings suggest that Italian chemical SMEs are more advanced in qualitative and narrative sustainability reporting than in systematic, quantitative, and VSME-specific disclosure.

Keywords: Voluntary sustainability reporting; Small and medium-sized enterprises (SMEs); European sustainability standards; Environmental sustainability; Chemical Sector

Paper type: Academic Research Paper

1 Introduction

The European Commission consistently highlights that the transition towards a more sustainable and digitally advanced economy depends on the effective involvement of small and medium-sized enterprises (SMEs), given their structural importance within the business landscape (Schulze et al., 2025).

Indeed, from a structural perspective, SMEs represent 99.8% of all enterprises in the EU non-financial business sector and contribute substantially to employment and value creation. They therefore constitute a cornerstone of the European economic system, accounting for a large share of employment creation and value added across ecosystems and, in several sectors, acting as a primary driver of net employment growth (Schulze et al., 2025).

Their capacity to engage in the green and digital transitions is therefore considered essential for achieving the objectives outlined in the European Green Deal, and SMEs will be required to transform both digitally and environmentally (European Commission, 2025a). This double challenge is likely to be particularly demanding for enterprises, especially considering the supply-chain pressures exerted both upstream and downstream. These pressures, both explicit and implicit, increasingly require SMEs to align with evolving expectations in terms of Corporate Social Responsibility (Harness et al., 2018; Choi et al., 2019; Stekelorum, Courrent, & Spence, 2024) and digital transformation (Mcilveene, 2026; Pálka et al., 2026). At the same time, SMEs face well-documented

structural constraints in engaging with digital and sustainability-related practices, including limited financial resources, skills shortages, and constrained managerial capacity (Álvarez Jaramillo, Zartha Sossa, & Orozco Mendoza, 2019; Madrid-Guijarro & Duréndez, 2024; Leipziger et al., 2025).

These challenges are particularly evident in the field of sustainability reporting, where disclosure frameworks have traditionally been developed with large and resource-rich firms in mind (Aureli et al., 2020; De Micco et al., 2021). With reference to sustainability reporting in SMEs, the literature consistently highlights structural limitations that affect both the adoption and the quality of disclosure practices. Resource constraints, in terms of time, financial capacity, and technical expertise, represent a primary barrier, limiting SMEs' ability to implement structured reporting processes and align with established standards. Consequently, reporting among SMEs tends to be predominantly qualitative and narrative-based, fragmented, weakly standardised, and often difficult to compare across firms (Steidle et al., 2025).

Within this context, the recent introduction of the Corporate Sustainability Reporting Directive (CSRD) and the European Sustainability Reporting Standards (ESRS) represents a major shift in the European sustainability reporting landscape (European Union, 2022; EFRAG, 2022). One of the main innovations of the CSRD is the extension of reporting obligations to listed SMEs, which represent only a small share of the overall SME population in Europe (European Commission, 2021). As a consequence, most SMEs remain outside the scope of mandatory reporting requirements. Nevertheless, many SMEs may still be indirectly affected through the so-called trickle-down effect, whereby firms subject to reporting obligations request sustainability-related information from companies within their value chains. This dynamic may generate additional reporting burdens for non-listed SMEs, even though they are not formally subject to sustainability reporting requirements (European Commission, 2025b).

In this respect, SMEs may face significant challenges in identifying, collecting, and managing sustainability-related data, both internally and along the value chain, particularly when dealing with the double materiality perspective and the interpretation of ESG reporting standards (Cavaliere, Bellazzecca, & di Castiglionchio, 2026). These structural and operational constraints are reflected in the still limited diffusion of environmental measurement practices. For instance, only a small proportion of SMEs have implemented advanced monitoring systems, such as the measurement of Scope 3 emissions or detailed carbon footprints (Schulze et al., 2025). This indicates a substantial gap between regulatory expectations and actual reporting capabilities, suggesting that the primary barrier to sustainability disclosure is not only reporting itself, but also the upstream process of data generation and measurement.

In response to concerns related to proportionality and administrative burden, the European Commission has promoted simplified and voluntary frameworks tailored to SMEs, most notably the Voluntary Sustainability Reporting Standard for non-listed SMEs (VSME) (EFRAG, 2023; Momtaz & Parra, 2025). As stated in the objectives of the VSME (EFRAG, 2024), the standard has been developed to support non-listed micro-, small- and medium-sized enterprises in addressing increasing sustainability-related information demands. In particular, it aims to enable SMEs to respond to data requests from large companies within the supply chain, as well as from banks and investors, thereby facilitating access to finance.

At the same time, the VSME is designed to support the internal management of environmental, social and governance issues, contributing to improved organisational practices, enhanced resilience, and long-term competitiveness. Although non-listed SMEs fall outside the scope of the CSRD, they are encouraged to adopt the VSME, which covers the same broad sustainability topics as the ESRS but in a more proportionate and simplified manner (EFRAG, 2024).

In this evolving regulatory context, assessing not only whether SMEs disclose sustainability information, but also the extent to which their voluntary disclosures are aligned with emerging European reporting standards becomes increasingly relevant. Prior research on sustainability reporting quality has shown that the presence of information alone is insufficient to ensure transparency and comparability, as firms may engage in symbolic disclosures that provide limited informational value (Michelon et al., 2015). However, empirical evidence on the quality and standard-alignment of sustainability reporting by SMEs remains scarce, particularly with regard to non-listed firms operating in environmentally intensive sectors and to the extent to which their sustainability practices are formally communicated through reporting mechanisms (Galli, Torelli, & Caccialanza, 2023).

This gap is especially relevant in the chemical industry. The chemical sector is among the largest manufacturing sectors in Europe, the third-largest manufacturing industry in the EU27, with over 30,000 SMEs accounting for a substantial share of employment and sales (Cefic, 2025). It is also one of the most environmentally relevant industries, in terms of greenhouse gas emissions, energy consumption, resource use, pollution, and potential impacts along the value chain. While sustainability reporting practices of large chemical corporations have been widely analysed, little is known about how chemical SMEs engage with voluntary sustainability reporting and how aligned their disclosures are with emerging European reporting standards.

Against this background, this paper assesses the quality of voluntary sustainability reporting by SMEs operating in the Italian chemical sector. More specifically, it evaluates the degree of alignment between firms' sustainability reports, or equivalent sustainability-related disclosures, and the EFRAG Voluntary Sustainability Reporting Standard for non-listed SMEs. In this study, reporting quality is operationalised as the degree of alignment with the VSME disclosure requirements. Empirically, the analysis focuses on firms classified under the C20 division of the NACE industrial classification, corresponding to the manufacture of chemicals and chemical products.

The study addresses the following research question:

RQ1. To what extent do the voluntary sustainability disclosures of Italian chemical SMEs align with the VSME requirements, and how does this alignment vary across modules, thematic areas and disclosure items?

By addressing this question, the paper contributes to the literature on SME sustainability reporting by providing empirical evidence on the degree of standard-alignment of voluntary disclosures in an environmentally relevant manufacturing sector. It also offers insights into the disclosure areas in which SMEs appear more advanced and those where alignment with the VSME remains more limited.

The remainder of the paper is structured as follows. The next section presents the research design, describing the sample selection procedure, the construction of the VSME coding grid, the scoring approach, and the quality-control process applied to document extraction. The following section reports the empirical results, examining VSME alignment at different levels of analysis, including score distribution, firm-level performance, modules, thematic areas, disclosure items, and clusters. Finally the paper then concludes by discussing the main findings, outlining the methodological limitations of the study, and identifying possible directions for future research.

2 Methodology

The empirical analysis was designed to assess the degree of alignment between the sustainability reports published by chemical Italian firms and the EFRAG Voluntary Sustainability Reporting Standard for non-listed SMEs (VSME).

The starting population was selected on the AIDA database, that is one of the main Italian databases providing corporate, financial, and economic information on firms. It was developed by Bureau van Dijk, now part of Moody's, and is widely used in academic research, economic analysis, due diligence activities, and business studies (Bureau van Dijk, 2024; Moody's, 2024). The initial sample consisted of 5,623 firms, of which 3,517 were identified as SMEs according to the European Commission classification based on the number of employees, annual turnover, and total balance sheet assets. Among these 3,517 firms, 109 were identified as having published a sustainability report or an equivalent sustainability-related disclosure document.

Given the high cost of manually collecting and coding corporate sustainability reports, a stratified sampling strategy was adopted. The sample was drawn from firms classified as having level-3 disclosure, using the relative distribution of the three clusters as the sampling criterion.

The final selected sample included 40 firms. For each firm, the relevant sustainability document was searched, downloaded where available, and stored in a structured folder. The documents included sustainability reports, impact reports, group or network sustainability reports, and sustainability-related corporat disclosures. The subsequent analysis was conducted through an AI-assisted content analysis protocol based on a VSME coding grid.

Table 1: Sampling design and final analytical dataset

Analytical step	N	Reference / percentage
Initial firm population	3517	100.00% of initial population
Firms with sustainability report / level-3 disclosure	109	3.10%
Stratified sample selected for document collection	40	36.70% of reporting firms
Reports successfully processed	33	82.50% of selected sample
Reports not included in final item-level analysis	7	17.50% of selected sample
Final firm-item observations	660	33 firms x 20 VSME items

The VSME coding grid was constructed by translating the standard into 20 disclosure items: 11 items from the Basic Module, labelled B1 to B11, and 9 items from the Comprehensive Module, labelled C1 to C9 (Table 2). Each item was assigned to one of four thematic areas: General information, Environment, Social, and Governance.

Table 2: VSME disclosure-item coding framework by module, sustainability area and coding requirement

ID Item	Module	Area	Item
B1	Basic	General information	Basis for preparation
B2	Basic	General information	Practices, policies and future initiatives for transitioning towards a more sustainable economy
B3	Basic	Environment	Energy and greenhouse gas emissions
B4	Basic	Environment	Pollution of air, water and soil
B5	Basic	Environment	Biodiversity
B6	Basic	Environment	Water
B7	Basic	Environment	Resource use, circular economy and waste management
B8	Basic	Social	Workforce - General characteristics
B9	Basic	Social	Workforce - Health and safety
B10	Basic	Social	Workforce - Remuneration, collective bargaining and training
B11	Basic	Governance	Convictions and fines for corruption and bribery
C1	Comprehensive	General information	Strategy: Business Model and Sustainability-Related Initiatives
C2	Comprehensive	General information	Description of practices, policies and future initiatives for transitioning towards a more sustainable economy
C3	Comprehensive	Environment	GHG reduction targets and climate transition
C4	Comprehensive	Environment	Climate risks
C5	Comprehensive	Social	Additional workforce characteristics
C6	Comprehensive	Social	Additional own workforce information - Human rights policies and processes
C7	Comprehensive	Social	Severe negative human rights incidents
C8	Comprehensive	Governance	Revenues from certain sectors and exclusion from EU reference benchmarks
C9	Comprehensive	Governance	Gender diversity ratio in the governance body

For each firm-item pair, the report was assessed using a three-level scoring scale: 0 indicated that the item was not covered or not aligned with the VSME requirement; 0.5 indicated partial coverage or partial alignment; and 1 indicated full coverage and substantive alignment with the VSME requirement.

As reported in Table 3, before analysing the results, a quality-control procedure was conducted on the document extraction process. Out of the 40 selected firms, 33 reports were successfully processed. Three documents were not found, two PDF files contained insufficient extractable text and were therefore likely to be scanned documents, and two reports generated extraction errors due to technical issues related to PDF encryption (Table 4). The final analytical dataset therefore consists of 33 firms and 660 firm-item observations, corresponding to 20 VSME items for each successfully processed firm.

Table 3: Quality-control results of the PDF extraction and processing phase

Status	N	Percentage of selected sample
Successfully processed	33	82.50%
PDF not found	3	7.50%
Insufficient extractable text / likely scanned PDF	2	5.00%
Extraction error	2	5.00%

Table 4: Firms excluded from the final item-level analysis

Firm name	Processing status	Reference year	Document type
AMBRO - SOL S.R.L. SOCIETA' BENEFIT , IN FORMA ABBREVIATA AMBRO - SOL S.R.L. SB	PDF not found	2024	Relazione di impatto società benefit
CROMATOS - S.R.L.	Insufficient extractable text / likely scanned PDF	2024	report ESG / bilancio di sostenibilità
EDILTECO - SOCIETA' PER AZIONI	Insufficient extractable text / likely scanned PDF	2023/2024	Bilancio di sostenibilità di gruppo
KULZER S.R.L.	Extraction error	2025	ESG report di gruppo
MERSEN ITALIA MALONNO S.R.L.	PDF not found	2024	Universal Registration Document / sustainability di gruppo

Firm name	Processing status	Reference year	Document type
SUNSTAR ITALIANA S.R.L.	PDF not found	2024	Sustainability Report di gruppo
SYMRISE S.R.L.	Extraction error	2024	Sustainability/corporate report di gruppo

3 Results Analysis

The distribution of item-level scores suggests a moderate degree of alignment with the VSME. As reported in Table 5, out of 660 firm-item observations, 141 observations received a score of 0, corresponding to 21.36% of the total; 256 observations received a score of 0.5, corresponding to 38.79%; and 263 observations received a score of 1, corresponding to 39.85%. This indicates that full alignment is observed in slightly less than 40% of the assessed disclosure items, while partial alignment is also similary frequent. Non-coverage remains substantial, but it is not the dominant outcome.

Table 5: Distribution of VSME item-level scores

Score	Interpretation	N	Percentage
0	Not covered / not aligned	141	21.36%
0.5	Partially covered / partially aligned	256	38.79%
1	Fully covered / aligned	263	39.85%

As summarized in Table 6, at the firm level, the average VSME alignment score is 0.592. The median score is 0.575, while the minimum and maximum values are 0.300 and 0.900, respectively. The standard deviation is 0.157. These results indicate a substantial degree of heterogeneity across firms in the extent to which their sustainability disclosures align with the VSME structure.

The proximity between the mean and the median suggests that the distribution of firm-level scores is not strongly driven by extreme observations. However, the relatively wide range between the lowest and highest values points to marked differences in disclosure practices within the sample. Some companies publish sustainability reports that cover only a limited subset of the VSME disclosure items, resulting in weak alignment with the standard. In these cases, sustainability information is likely to be present but fragmented, selective, or not organised according to the categories and information requirements reflected in the VSME framework.

Conversely, the firms with higher alignment scores disclose a broader and more structured set of sustainability information. Their reports appear to cover a larger number of VSME-related items and to provide information across multiple areas. Overall, the evidence suggests that, even among firms that publish a sustainability report or an equivalent sustainability-related disclosure, the degree of consistency with the VSME framework varies considerably. This confirms that the mere presence of a sustainability document does not necessarily imply comprehensive or standard-aligned reporting.

Table 6: Descriptive statistics of firm-level VSME alignment scores

Statistic	Value
Number of firms	33
Mean	0.592
Median	0.575
Minimum	0.300
25th percentile	0.500
75th percentile	0.700
Maximum	0.900
Standard deviation	0.157

The comparison between the Basic and Comprehensive Modules shows relatively similar average scores. The Basic Module has an average score of 0.588, while the Comprehensive Module has an average score of 0.598 (Table 7). This result is noteworthy because the Comprehensive Module contains more demanding disclosures, especially on climate risks, GHG targets, human rights incidents, governance diversity, and revenues from sensitive sectors. The slightly higher score for the Comprehensive Module is mainly driven by strong performance on business model, sustainability initiatives, and human-rights-policy disclosures.

Table 7: VSME alignment by module

Module	Mean score	Minimum	Maximum	Standard deviation	N firm-item observations
Basic	0.588	0.000	1.000	0.343	363
Comprehensive	0.598	0.000	1.000	0.422	297

The analysis by thematic area reveals clearer differences, as shown in Figure 1 and summarized in Table 8. General information is the most aligned area, with an average score of 0.814. This suggests that firms tend to disclose basic contextual information, business-model elements, and descriptions of sustainability initiatives relatively well. The Social area also shows a comparatively strong level of alignment, with an average score of 0.644. Environment follows with an average score of 0.511, indicating intermediate alignment. Governance is the weakest area, with an average score of 0.384. This result suggests that governance-related information, especially corruption and bribery sanctions, sensitive-sector revenues, benchmark exclusion, and board gender diversity, remains less systematically reported.

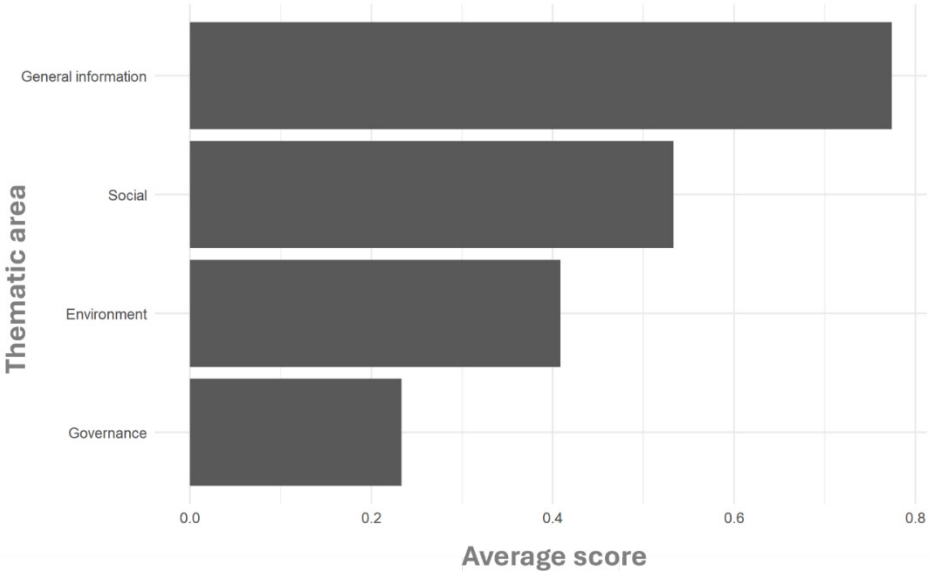


Figure 1: Average VSME alignment score by thematic area

Table 8: VSME alignment by thematic area

Area	Mean score	Minimum	Maximum	Standard deviation	N firm-item observations
General information	0.814	0.500	1.000	0.242	132
Social	0.644	0.000	1.000	0.346	198
Environment	0.511	0.000	1.000	0.362	231
Governance	0.384	0.000	1.000	0.462	99

At the item level, the strongest alignment is observed for C1, Strategy: Business Model and Sustainability-Related Initiatives, with an average score of 1.000. This means that all successfully analysed reports provide sufficient information on the firm's business model and strategic sustainability-related initiatives. High scores are also observed for C2, Description of practices, policies and future initiatives, with an average of 0.894; B2, Practices, policies and future initiatives for transitioning towards a more sustainable economy, with an average of 0.864; C6, Human rights policies and processes, with an average of 0.864; and B7, Resource use, circular economy and waste management, with an average of 0.803.

The weakest items are C8, Revenues from certain sectors and exclusion from EU reference benchmarks, with an average score of 0.121; B5, Biodiversity, with an average score of 0.212; B4, Pollution of air, water and soil, with an average score of 0.333; B11, Convictions and fines for corruption and bribery, with an average score of 0.409; and C3, GHG reduction targets and climate transition, with an average score of 0.439. These results suggest that firms are more likely to provide narrative descriptions of sustainability practices, business models, and workforce-related policies than detailed quantitative disclosures on environmental impacts, governance incidents, or climate-transition targets.

Cluster-level results should be interpreted with caution. Although the sampling strategy was designed to reflect the original cluster structure, the final successfully processed reports are concentrated in cluster 2, with

31 firms, and cluster 3, with only 2 firms. No firm from cluster 1 appears in the final processed item-level dataset. The average score for cluster 2 is 0.579, while the average score for cluster 3 is 0.800. However, given the very small number of observations in cluster 3, this comparison should be treated as descriptive only and not as evidence of statistically robust differences across clusters.

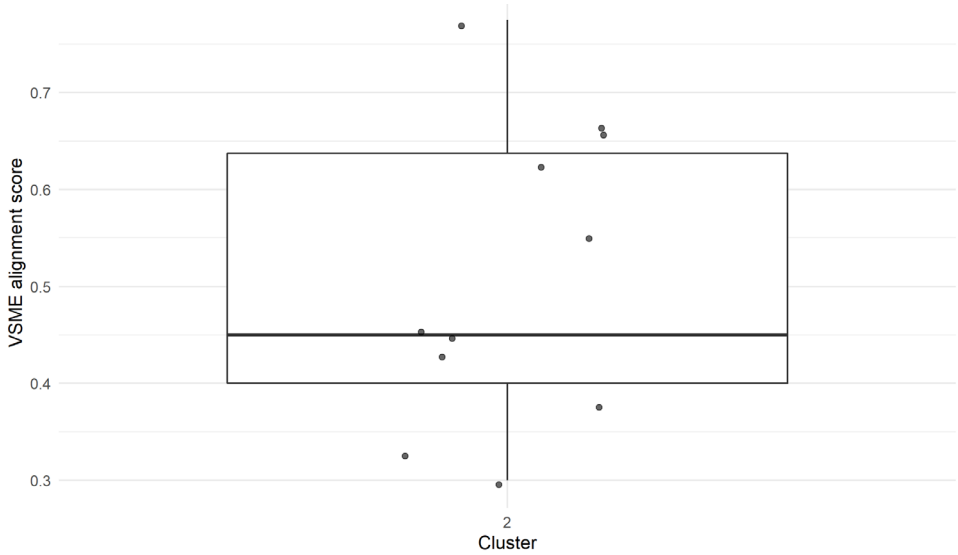


Figure2: Average VSME alignment by cluster

4 Conclusions

Overall, the results show that the analysed Italian firms disclose sustainability information in a way that is partially aligned with the VSME, but alignment is uneven across areas and items. The strongest disclosure practices concern general information, business-model descriptions, sustainability initiatives, and social policies. Conversely, the weakest areas concern governance disclosures, biodiversity, pollution, climate-transition targets, and sensitive-sector revenue information. This pattern suggests that firms are more advanced in qualitative and narrative sustainability reporting than in systematic, quantitative, and VSME-specific disclosure.

From a methodological perspective, the AI-assisted content analysis allowed the systematic coding of a relatively large number of disclosure items across multiple corporate reports. However, the analysis also highlights the importance of document availability and text-extraction quality. Seven out of the 40 selected firms could not be included in the final item-level analysis because the report was missing, technically unreadable, or insufficiently extractable. Future refinements of the analysis may therefore include manual recovery of missing or scanned reports, OCR processing for image-based PDFs, and manual validation of a subset of AI-coded observations to assess coding reliability.

5 Ethics Declaration

This study did not involve human participants, interviews, surveys, experiments, or the collection of personal data. The empirical analysis was based exclusively on publicly available corporate sustainability reports and equivalent sustainability-related disclosure documents. For this reason, no formal ethical clearance was required. All documents were analysed for academic research purposes and are reported only in aggregated form.

6 AI Declaration

This study used AI-assisted tools to support the content analysis of corporate sustainability reports. The AI tool was employed to assist in identifying and coding disclosure information according to the VSME coding grid developed for the study.

AI-assisted tools were also used during the writing process to improve the clarity, readability, grammar, and academic style of selected parts of the manuscript. The use of these tools did not replace the authors' judgement in the design of the study, the definition of the coding framework, the interpretation of the results, or the development of the conclusions.

After using these tools, the authors reviewed, edited, and validated the generated outputs and take full responsibility for the content of the submitted manuscript.

References

Álvarez Jaramillo, J., Zartha Sossa, J. W., & Orozco Mendoza, G. L. (2019). Barriers to sustainability for small and medium enterprises in the framework of sustainable development—Literature review. *Business Strategy and the Environment*, *28*(4), 512-524.

Aureli, S., Del Baldo, M., Lombardi, R., & Nappo, F. (2020). Nonfinancial reporting regulation and challenges in sustainability disclosure and corporate governance practices. Business Strategy and the Environment, 29(6), 2392–2403.

Cavaliere, F., Bellazzecca, E., & di Castiglionchio, E. E. Z. (2026). Walking the Tightrope of Sustainability Reporting Standardization: Insights From Italian Small and Medium-Sized Enterprises. *Corporate Social Responsibility and Environmental Management*.

CEFIC (2025) The European Chemical Industry – A Vital Part of Europe's Future: Facts and Figures, The European Chemical Industry Council, Brussels, Belgium, [online], available at: https://cefic.org/a-pillar-of-the-european-economy/facts-and-figures-of-the-european-chemical-industry/ (accessed 22 January 2026).

Choi, S. B., Feng, Y., Liu, J., & Zhu, Q. (2019). Motivating corporate social responsibility practices under customer pressure among small-and medium-sized suppliers in China: The role of dynamic capabilities. *Corporate Social Responsibility and Environmental Management*, *26*(1), 213-226.

De Micco, P., Rinaldi, L., Vitale, G., Cupertino, S., & Maraghini, M. P. (2021). The challenges of sustainability reporting and their management: The case of Estra. Meditari Accountancy Research, 29(3), 430–448.

EFRAG. (2022). European Sustainability Reporting Standards (ESRS): General Requirements and Environmental Standards. European Financial Reporting Advisory Group, Brussels.

EFRAG. (2023). Voluntary Sustainability Reporting Standard for SMEs (VSME): Exposure Draft. European Financial Reporting Advisory Group, Brussels.

EFRAG. (2024). *Voluntary sustainability reporting standard for non-listed micro-, small- and medium-sized undertakings (VSME)*. European Financial Reporting Advisory Group.

European Commission. (2021). *Commission staff working document: Impact assessment accompanying the document Proposal for a Directive of the European Parliament and of the Council amending Directive 2013/34/EU, Directive 2004/109/EC, Directive 2006/43/EC and Regulation (EU) No 537/2014, as regards corporate sustainability reporting* (SWD/2021/150 final)

European Commission, Joint Research Centre, Di Bella, L., Lagüera González, J., Katsinis, A., & Dominguez Torreiro, M. (2025a). The SME twin transition monitor for the EU (JRC137875). Publications Office of the European Union.

European Commission. (2025b). *Commission Recommendation of 30.7.2025 on a voluntary sustainability reporting standard for small and medium-sized undertakings* (C/2025/4984 final).

European Union. (2022). Directive (EU) 2022/2464 of the European Parliament and of the Council on corporate sustainability reporting (CSRD). Official Journal of the European Union.

Galli, D., Torelli, R., & Caccialanza, A. (2023). Sustainability performance and sustainability reporting in SMEs: A love affair or a fight? Journal of Management & Organization, 29(6), 1069–1090. https://doi.org/10.1017/jmo.2023.40

Harness, D., Ranaweera, C., Karjaluoto, H., & Jayawardhena, C. (2018). The role of negative and positive forms of power in supporting CSR alignment and commitment between large firms and SMEs. *Industrial Marketing Management*, *75*, 17-30.

Leipziger, M., Burtscher, J., Kanbach, D. K., Puumalainen, K., & Kraus, S. (2025). WE HAVE ALWAYS DONE IT THAT WAY—BARRIERS AND ENABLERS FOR THE TWIN TRANSFORMATION OF SME s. *International journal of innovation management*, *29*(05n06), 2540004.

Madrid-Guijarro, A., & Duréndez, A. (2024). Sustainable development barriers and pressures in SMEs: The mediating effect of management commitment to environmental practices. *Business Strategy and the Environment*, *33*(2), 949-967.

Mcilveene, T. R. (2026). Advancing Small Business Strategy in Turbulent Times: Capabilities, Ethical AI, Customer Retention, Brand Trust, Workplace Culture, and Trade Policy Uncertainty. *Journal of Small Business Strategy*, *36*(2), 1-6.

Michelon, G., Pilonato, S., & Ricceri, F. (2015). CSR reporting practices and the quality of disclosure: An empirical analysis. Critical Perspectives on Accounting, 33, 59–78. https://doi.org/10.1016/j.cpa.2014.10.003

Momtaz, P. P., & Parra, I. M. (2025). Is sustainable entrepreneurship profitable? ESG disclosure and the financial performance of SMEs. *Small Business Economics*, *64*(4), 1535-1564.

Pálka, P., Blahová, M., Ntsiful, A., & Kwarteng, M. A. (2026). Balancing external pressure and risk tolerance: examining the influence of business value in entrepreneurs' investment decisions in digital technologies. *International Entrepreneurship and Management Journal*, *22*(2), 77.

Schulze, B.P., Katsinis, A., Lagüera, G.J., Di Bella, L., Odenthal, L., Hell, M., … Secades, C.B. (2025) Annual Report on European SMEs 2024/2025: SME Performance Review, Publications Office of the European Union, Luxembourg.

Steidle, C., Ostojic, S., Achterfeldt, S., & Traverso, M. (2025). Small players, big impact? Unveiling practices and challenges of sustainability reporting by German SMEs. *Discover Sustainability*, *6*(1), 780.

Stekelorum, R., Courrent, J. M., & Spence, M. (2024). Supply chain pressures and SMEs' CSR practices: the moderation effect of supply chain position. *International Journal of Logistics Research and Applications*, *27*(11), 2127-2152.

Driving Innovation and Digital Transformation in Cultural Tourism Organizations and Local Ecosystems: Final Insights of the T4Y Project

Francesco Santarsiero, Daniela Carlucci, Antonio Lerro
University of Basilicata, Potenza, Italy

Abstract

Digital transformation is increasingly presented as a necessary pathway for cultural tourism organizations, yet many local and rural ecosystems still struggle to translate digital opportunities into coherent strategies, organizational capabilities, and sustainable business model innovation. This paper presents the final insights of the T4Y research project, which developed and applied models and tools for analysing, mapping and supporting digital transformation and business model innovation in cultural tourism organizations and local ecosystems. The study is grounded in the Tech4You project and adopts an iterative, multi-method and design-oriented research approach. The research process combined an international literature review on digital transformation, digital maturity and business model innovation; an analysis of best practices; the development of a sector-specific Digital Maturity Assessment Model and related assessment tool; qualitative interviews with local stakeholders, operators and institutional actors; survey-based data collection; and participatory workshops in pilot areas in Basilicata, Southern Italy. Findings show that digital transformation in cultural tourism cannot be interpreted as a mere process of technology adoption. Rather, it emerges as a systemic capability involving seven interconnected dimensions: strategy, organization, culture, employees, tourists/customers, operations and technology. The final project insights highlight four main issues. First, digital maturity is uneven and often limited by weak strategic planning, fragmented technological infrastructures, limited use of data and discontinuous collaboration among actors. Second, business model innovation is activated when digital tools are connected to value proposition redesign, stakeholder engagement, experience innovation and territorial identity. Third, participatory tools such as assessment models, SWOT-based reflection, business model mapping and workshops act as boundary objects that transform scattered knowledge into actionable ecosystem learning. Fourth, leadership and relational capabilities are decisive for moving from isolated digital initiatives to coordinated transformation pathways. The paper contributes to knowledge management, digital transformation and tourism innovation research by framing digital maturity as an ecosystem-level knowledge capability and by offering a practical roadmap for cultural tourism organizations and policy makers.

Keywords: Cultural tourism; Digital transformation; Digital maturity; Business model innovation; Innovation ecosystems.

Paper type: Academic Research Paper

1 Introduction

Cultural tourism organizations are facing a double transformation. On the one hand, they are required to preserve, interpret and communicate cultural, historical and natural heritage in ways that remain faithful to local identity. On the other hand, they must respond to new visitor behaviours, platform-based intermediation, data-driven decision-making, sustainability pressures and the growing expectation that tourism experiences should be accessible, personalized and digitally supported. These pressures are particularly relevant in rural and peripheral territories, where cultural and natural resources can represent strong development assets, but where organizations are often small, fragmented and characterized by limited financial and managerial resources.

The debate on digital transformation has progressively moved beyond a narrow understanding of technology adoption. Digital transformation involves changes in strategy, processes, capabilities, culture, value creation logics and inter-organizational relationships (Vial, 2019; Verhoef et al., 2021). In tourism, this change is strongly connected to the smart tourism literature, which emphasizes the integration of digital infrastructures, data, stakeholders and visitor experience (Gretzel et al., 2015; Buhalis and Amaranggana, 2014). However, in cultural tourism ecosystems the adoption of digital tools does not automatically generate innovation. Digital technologies become valuable when they are embedded in organizational learning processes, shared visions and collaborative mechanisms capable of connecting individual organizations with wider territorial systems.

This paper presents the final insights of the T4Y research project, developed within the Tech4You ecosystem, with specific reference to the analysis, mapping and assessment of digital transformation and business model innovation in cultural tourism organizations. The project was designed to support local ecosystems in Basilicata, Southern Italy, by developing models and tools that could combine scientific robustness with practical usability.

Rather than producing a generic digital checklist, the project aimed to understand how cultural tourism organizations can build digital maturity, redesign their business models and activate collaborative innovation dynamics at the ecosystem level.

The paper is positioned as a synthesis and conceptual integration of the main research activities carried out during the project. It builds on previous outputs related to the Digital Maturity Assessment Model, participatory events, digital leadership and business model innovation tools, and it offers a final interpretation of what the project reveals about the digital transformation of cultural tourism ecosystems. The central research question is: How can digital transformation and business model innovation be assessed and supported in cultural tourism organizations and local ecosystems?

To address this question, the paper adopts a multi-method and design-oriented perspective. It reconstructs the development and application of a toolkit composed of a Digital Maturity Assessment Model, a survey-based assessment tool, qualitative interviews, participatory workshops, SWOT-based reflection, business model mapping and a digital transformation roadmap. The contribution is twofold. Theoretically, the paper frames digital maturity as a knowledge-based and ecosystem-level capability, rather than as an internal technological condition of individual organizations. Practically, it offers actionable insights for managers, policy makers and territorial innovation actors interested in strengthening the digital readiness and innovative capacity of cultural tourism systems.

2 Theoretical Background

2.1 Digital transformation and digital maturity in cultural tourism

Digital transformation can be understood as a process through which digital technologies trigger substantial changes in organizational strategy, structures, processes and value creation mechanisms (Bharadwaj et al., 2013; Vial, 2019). This definition is particularly important for tourism and cultural organizations because it avoids reducing digital transformation to the adoption of websites, social media profiles or isolated technological applications. The issue is not only whether an organization uses digital tools, but whether it is able to use them to rethink its relationship with visitors, partners, knowledge resources and territorial value.

Digital maturity models provide a useful way to assess this readiness. They help organizations understand their current position, identify capability gaps and define transformation priorities. However, many maturity models remain either generic or strongly technology-centred. In cultural tourism, this limitation is problematic because digitalization is closely linked to intangible resources, local knowledge, cultural narratives, stakeholder relationships and experience design. A museum, a village, a natural park, an accommodation provider or a destination management actor may all need digital capabilities, but these capabilities must be interpreted in relation to different missions, audiences and territorial roles.

For this reason, the T4Y project developed a sector-specific Digital Maturity Assessment Model structured around seven dimensions: strategy, organization, culture, employees, tourists/customers, operations and technology. These dimensions were conceived as interdependent. Strategy captures the degree to which digital transformation is connected to vision, positioning, collaboration and innovation priorities. Organization concerns structures, learning mechanisms and change management. Culture refers to digital mindset, openness to innovation and entrepreneurial orientation. Employees concern awareness, skills and continuous learning. Tourists/customers focus on digital experience, engagement and relationship management. Operations address data, analytics, resource management and process integration. Technology refers to information systems, infrastructure, security and privacy. Together, these dimensions provide an integrated perspective on digital maturity.

2.2 Business model innovation and ecosystem knowledge

Business model innovation refers to changes in the way organizations create, deliver and capture value (Teece, 2010; Foss and Saebi, 2017). In cultural tourism, business model innovation is not limited to revenue models or digital channels. It includes the redesign of value propositions, visitor journeys, partnerships, knowledge flows, cultural narratives and social impact mechanisms. Digital technologies can support these transformations, but only when they are connected to a broader understanding of value.

Cultural tourism ecosystems are characterized by interdependence. Visitors do not experience a single organization in isolation; they encounter a combination of heritage, hospitality, mobility, food, events, landscape, digital information, storytelling and community relations. Therefore, the innovation capacity of a single actor is often constrained or enabled by the maturity of the wider ecosystem. This perspective is

consistent with ecosystem theory, which highlights the role of complementary actors, shared structures and value co-creation (Adner, 2017; Jacobides et al., 2018). It also resonates with knowledge management research, which emphasizes knowledge creation, sharing and application as collective processes (Nonaka and Takeuchi, 1995; Alavi and Leidner, 2001).

From this viewpoint, digital maturity can be interpreted as a distributed knowledge capability. It depends on the ability of local actors to collect and interpret information, share experiences, align strategic priorities, learn from experimentation and transform tacit territorial knowledge into accessible and sustainable visitor experiences. The T4Y project therefore treated assessment tools and participatory workshops not merely as data collection instruments, but as knowledge translation mechanisms able to make digital transformation visible, discussable and actionable for local actors.

3 Methodology

The study is based on the research activities carried out within the Tech4You project, with a specific focus on cultural tourism organizations and local ecosystems in Basilicata. The research design was iterative, multi-method and design-oriented. Its aim was not only to observe digital transformation, but also to build, test and refine models and tools able to support organizations and territorial actors in their transformation pathways.

The first phase consisted of a literature review on digital transformation, digital maturity, business model innovation, tourism innovation and cultural tourism ecosystems. This phase supported the identification of the main theoretical constructs and informed the design of the assessment model. In parallel, an analysis of national and international best practices was carried out to understand how cultural and tourism organizations approach digital technologies, stakeholder collaboration, visitor engagement and business model redesign.

The second phase focused on tool development. The research team developed a Digital Maturity Assessment Model and a survey-based assessment tool. The tool was then refined through feedback and pilot testing, with the aim of improving clarity, internal coherence and applicability to different types of cultural tourism actors. Additional methodological tools were developed or adapted to connect digital maturity assessment with business model innovation. These included a Business Model Prism for cultural tourism, SWOT-based reflection, digital leadership prompts and workshop formats aimed at stimulating participatory discussion.

The empirical phase combined quantitative and qualitative data. Survey-based assessment was used to collect structured and comparable information on the maturity dimensions. Qualitative interviews with tourism operators, cultural organizations, institutional representatives and local stakeholders were used to interpret the reasons behind the levels of maturity observed and to explore decision-making processes, organizational capabilities, change management practices, collaboration mechanisms and leadership attitudes. Participatory workshops were organized as spaces for validation, knowledge sharing and co-creation. They allowed local actors to discuss results, identify needs and opportunities, and translate evidence into possible innovation directions.

The project covered five pilot areas, enabling a comparative reading of digital maturity and business model innovation dynamics across different territorial contexts. The analysis was conducted through descriptive synthesis, cross-case comparison and thematic interpretation. Rather than treating the data as a purely statistical exercise, the analysis aimed to identify recurring patterns and actionable insights emerging from the combination of survey results, interviews and participatory interactions. This approach is coherent with the practical and theoretical ambition of the project: to produce models and tools that are grounded in empirical evidence while remaining usable for organizations and policy makers.

Table 1: The Digital Maturity Assessment Model developed in the T4Y project

Dimension	Assessment focus	Interpretative role
Strategy	Digital vision, positioning, communication, collaboration and innovation priorities.	Clarifies whether digital transformation is part of a coherent strategic agenda rather than a set of isolated initiatives.
Organization	Structures, roles, learning mechanisms, change management and coordination processes.	Shows whether the organization can translate digital intentions into routines and responsibilities.
Culture	Digital mindset, openness to innovation, experimentation and entrepreneurial attitude.	Captures the cultural conditions that support or constrain transformation.
Employees	Digital awareness, skills, continuous learning and capability development.	Highlights the human capability base needed to use technologies effectively.

Dimension	Assessment focus	Interpretative role
Tourists/customers	Digital experience, engagement, relationship management and visitor understanding.	Connects digital maturity to experience quality and value co-creation.
Operations	Data use, analytics, KPIs, process integration, automation and resource management.	Indicates whether digital tools support operational learning and evidence-based decisions.
Technology	Information systems, infrastructure, interoperability, security and privacy.	Assesses the enabling technological foundations of transformation.

4 Findings: Final Insights from the T4Y Project

4.1 Digital maturity is multidimensional and uneven

The first final insight is that digital maturity in cultural tourism organizations is multidimensional and uneven. The project confirmed that organizations may be relatively active in some visible digital practices, such as online communication or social media use, while remaining weak in strategic planning, operational data use, integrated systems or capability development. This means that digital presence should not be confused with digital maturity. A cultural tourism organization may be visible online without having a clear digital strategy, structured visitor data, integrated processes or a shared understanding of how digital tools contribute to long-term value creation.

Across the pilot areas, the most critical issues concerned the fragmentation of digital initiatives, the limited formalization of digital objectives and the difficulty of connecting technological choices to measurable outcomes. Many actors recognized the importance of digital transformation, but this recognition did not always translate into systematic planning or investment priorities. The operational dimension was particularly relevant: data, analytics and key performance indicators were often underused, limiting the capacity of organizations to learn from visitors, monitor processes and evaluate the impact of innovation initiatives.

At the same time, the assessment revealed positive conditions on which transformation can build. Cultural tourism actors possess strong territorial knowledge, relational proximity, authenticity and awareness of the value of heritage. These resources represent a strategic basis for digital transformation, provided that they are connected to appropriate capabilities and collaborative mechanisms. Digital transformation should therefore start from existing local strengths rather than importing generic technological solutions disconnected from place-based identity.

4.2 Digital transformation becomes valuable when connected to business model innovation

The second insight concerns the relationship between digital transformation and business model innovation. The project showed that digital tools generate limited value when they are used only to improve communication or automate existing activities. Their transformative potential emerges when they support a rethinking of the organization's value proposition, visitor relationships, partnerships, internal processes and social or cultural impact.

The Business Model Prism used in the project helped actors reflect on how cultural tourism organizations create and capture value. It encouraged them to move beyond the question of which technologies to adopt and to ask why digitalization is needed, for whom value is created, which resources and partnerships are required, and how benefits can be distributed across the ecosystem. For example, digital storytelling, immersive experiences, booking platforms, data dashboards, customer relationship management tools or collaborative online maps can all be relevant. However, their value depends on whether they contribute to clearer positioning, stronger engagement, better coordination or more sustainable use of cultural and natural resources.

This finding is important because it positions business model innovation as the bridge between technological adoption and ecosystem development. In cultural tourism, the business model is often hybrid: it includes economic, cultural, social and territorial forms of value. A digital transformation roadmap should therefore avoid a narrow efficiency logic and should include questions of accessibility, inclusion, heritage interpretation, community participation and long-term sustainability.

4.3 Participatory tools act as boundary objects for ecosystem learning

The third insight concerns the role of tools. The assessment model, survey, SWOT analysis, Business Model Prism and workshop formats did not simply produce data for researchers. They functioned as boundary objects: practical artefacts that enabled different actors to discuss digital transformation using a common language. This

was particularly important in fragmented ecosystems, where organizations, institutions and stakeholders may have different priorities, vocabularies and expectations.

Participatory workshops created spaces in which local actors could interpret assessment results, compare perspectives and identify shared priorities. In this sense, the project generated not only analytical knowledge, but also relational and actionable knowledge. The workshops helped transform digital maturity from an abstract concept into a concrete topic of discussion. They also supported the emergence of needs, barriers and opportunities that would be difficult to capture through survey data alone.

This finding has implications for knowledge management. In cultural tourism ecosystems, knowledge is distributed among many actors and is often tacit, experiential and place-based. Assessment tools become useful when they help externalize this knowledge, connect it with structured evidence and translate it into decisions. The value of a maturity model is therefore not only diagnostic. It is also educational and mobilizing, because it can stimulate awareness, align actors and support the design of coordinated actions.

4.4 Leadership and relational capabilities enable the transition from projects to transformation

The fourth insight is that leadership is decisive for moving from individual digital projects to systemic transformation. Interviews and stakeholder interactions highlighted that digital transformation depends on people who are able to interpret opportunities, mobilize resources, build trust and connect actors. In many cultural tourism organizations, leadership is not only a formal managerial function. It is also a relational capability distributed across entrepreneurs, public administrators, cultural operators, associations, technology providers and community representatives.

The T4Y project highlighted the relevance of pragmatic, sustainable, facilitative, communicative, catalytic and informed leadership attitudes. Pragmatic leadership helps translate broad digital ambitions into feasible actions. Sustainable leadership connects innovation with long-term territorial value. Facilitative leadership enables participation and collaboration. Communicative leadership makes digital transformation understandable. Catalytic leadership activates change and experimentation. Informed leadership supports decisions based on evidence rather than intuition alone.

These leadership capabilities are particularly important in local ecosystems, where no single actor controls the whole visitor experience or the entire transformation process. The challenge is therefore not only to improve the digital skills of individual organizations, but to develop the capacity of the ecosystem to sense opportunities, share knowledge, coordinate initiatives and scale successful practices.

4.5 The roadmap should be progressive, situated and ecosystem-oriented

The final insight concerns the roadmap for digital transformation. The project suggests that cultural tourism organizations and local ecosystems should avoid both technological acceleration without readiness and excessive planning without experimentation. A useful roadmap should be progressive, situated and ecosystem-oriented. It should start from awareness and assessment, move towards capability building and targeted experimentation, and then consolidate practices through governance, data use and collaboration.

A first step is to map digital maturity and identify shared priorities. A second step is to strengthen basic capabilities, including digital skills, data literacy, communication practices and awareness of available technologies. A third step is to redesign selected visitor experiences and internal processes, connecting technology to business model innovation. A fourth step is to build ecosystem infrastructures, such as common data practices, shared storytelling strategies, partnership mechanisms and innovation labs or workshops. A final step is to institutionalize learning through monitoring, evaluation, governance and continuous improvement.

This roadmap is not intended as a linear recipe. Cultural tourism ecosystems differ in resources, governance arrangements, digital readiness and stakeholder relationships. The roadmap should therefore be adapted to each local context. Its main value lies in helping actors move from scattered initiatives to coherent transformation pathways.

Table 2: Final insights and implications

Final insight	Ecosystem implication	Managerial and policy implication
Digital maturity is multidimensional and uneven.	Ecosystems may show digital visibility without strategic, operational or technological integration.	Assess maturity across strategy, organization, culture, employees, tourists/customers, operations and technology before defining interventions.

Final insight	Ecosystem implication	Managerial and policy implication
Digital transformation becomes valuable when linked to business model innovation.	Technology adoption should support value proposition redesign, experience innovation and territorial value creation.	Use business model mapping to connect digital tools with value, partners, processes and impact.
Participatory tools act as boundary objects.	Shared tools help fragmented actors create a common language and transform tacit knowledge into actionable priorities.	Combine surveys, interviews, workshops and SWOT-based reflection to support collective learning.
Leadership and relational capabilities enable transformation.	No single actor can lead the whole ecosystem; transformation requires distributed leadership and trust.	Develop facilitative, communicative, pragmatic, sustainable, catalytic and informed leadership capabilities.
Roadmaps should be progressive and situated.	Transformation depends on local identity, resources and collaboration maturity.	Move from assessment to capability building, experimentation, ecosystem infrastructure and continuous evaluation.

5 Discussion

The findings offer three main contributions to the literature on knowledge management, digital transformation and cultural tourism innovation. First, the paper advances the idea that digital maturity should be understood as an ecosystem-level knowledge capability. Existing approaches often focus on the digital readiness of individual organizations. The T4Y project shows that, in cultural tourism, the transformation capacity of a single organization is deeply connected to the knowledge, resources and coordination mechanisms of the wider ecosystem. Digital maturity is therefore not only a property of internal systems; it is also a relational condition that emerges through interaction among actors.

Second, the paper contributes to the debate on business model innovation by showing that cultural tourism organizations require value frameworks that include economic, cultural, social and territorial dimensions. Digital technologies can support new forms of value creation, but they must be interpreted through the specific mission of cultural tourism: preserving heritage, enriching visitor experience, involving communities and strengthening the attractiveness and resilience of territories. This perspective extends business model innovation research by emphasizing hybrid value creation and place-based knowledge.

Third, the paper contributes methodologically by showing how assessment tools can be used not only for measurement, but also for engagement and learning. The T4Y toolkit created a bridge between academic research and territorial practice. It allowed the research team to collect structured data, while also enabling local actors to understand, discuss and use evidence. This dual function is particularly relevant for IFKAD and knowledge management scholarship because it highlights how knowledge assets can be translated into innovation dynamics through participatory and actionable devices.

The paper also contributes to the debate on the theme of intelligent knowledge for sustainable organizations. Intelligent knowledge is not limited to advanced technological systems. In local cultural tourism ecosystems, it also concerns the capacity to combine data, human judgment, cultural identity, stakeholder experience and strategic action. The T4Y project suggests that sustainable digital transformation requires hybrid knowledge processes: structured assessment, participatory interpretation, leadership, experimentation and continuous learning.

6 Practical Implications

For managers of cultural tourism organizations, the main implication is that digital transformation should be governed as an organizational and strategic process. Before investing in technologies, organizations should clarify their digital objectives, assess their capabilities and identify how digital tools can improve visitor experience, internal processes and stakeholder relationships. They should also avoid treating communication channels as the whole of digital transformation. Social media presence, websites and online promotion are important, but they represent only one part of a broader transformation journey.

For policy makers and territorial development agencies, the findings suggest the importance of ecosystem-level interventions. Many small organizations cannot develop advanced digital capabilities alone. They need shared infrastructures, training opportunities, advisory services, data practices, innovation spaces and collaborative projects. Public policies should therefore support not only funding for technologies, but also capability building, knowledge sharing and governance mechanisms that connect cultural, tourism, technological and institutional actors.

For universities and research organizations, the project demonstrates the value of engaged research. The role of the university is not only to study digital transformation from outside, but also to design tools, facilitate workshops and support evidence-based decision-making. This approach strengthens the third mission of universities and contributes to the development of territorial innovation ecosystems.

7 Conclusions

This paper presented the final insights of the T4Y project on innovation and digital transformation in cultural tourism organizations and local ecosystems. The project developed and applied a set of models and tools aimed at assessing digital maturity, supporting business model innovation and activating stakeholder engagement in pilot areas in Basilicata. The results show that digital transformation in cultural tourism is a complex and multidimensional process that cannot be reduced to technology adoption.

The main conclusion is that cultural tourism ecosystems need to develop digital maturity as a collective capability. This capability includes strategy, organization, culture, employees, tourists/customers, operations and technology, but it also requires leadership, collaboration and knowledge translation. The T4Y project showed that assessment models and participatory tools can help actors make sense of their current condition, identify gaps and design more coherent transformation pathways.

The study has limitations. The empirical work is grounded in specific local contexts in Basilicata, and the findings should be further validated in other regions and cultural tourism ecosystems. Moreover, the paper provides an integrative synthesis of project results rather than a detailed statistical analysis of each pilot area. Future research could extend the assessment model to additional contexts, develop longitudinal analyses of digital maturity evolution, integrate the tools into digital platforms or decision-support systems, and examine how artificial intelligence can support data-driven and human-centred cultural tourism innovation.

Despite these limitations, the paper offers a practical and theoretical contribution. It provides a final synthesis of the T4Y project and shows how digital transformation, business model innovation and knowledge management can be connected to support more sustainable, adaptive and collaborative cultural tourism ecosystems.

8 Acknowledgements

This paper is based on the research activities carried out within the Tech4You project - Technologies for Climate Change Adaptation and Quality of Life Improvement, code ECS_00000009, CUP C43C22000400006, with specific reference to the activities related to digital transformation and business model innovation in cultural tourism organizations and ecosystems. The authors acknowledge the contribution of all organizations, stakeholders and institutional actors involved in the project activities and participatory workshops.

9 Ethics declaration

Ethical clearance was not required for this research according to the institutional procedures applicable to the project activities. The study involved organizational and stakeholder-level data collection, interviews and participatory workshops. No sensitive personal data are reported in this paper. Participation in interviews and workshops was voluntary and the results are presented in aggregated and interpretative form.

10 AI declaration

AI tools were used to support the editorial refinement of this manuscript. The authors take full responsibility for the accuracy, validity and integrity of the research design, data collection, analysis, interpretation and final manuscript. All AI-assisted text was critically reviewed, verified and accepted by the authors.

References

Adner, R. (2017), "Ecosystem as structure: An actionable construct for strategy", Journal of Management, Vol. 43 No. 1, pp. 39-58.

Alavi, M. and Leidner, D.E. (2001), "Review: Knowledge management and knowledge management systems: Conceptual foundations and research issues", MIS Quarterly, Vol. 25 No. 1, pp. 107-136.

Bharadwaj, A., El Sawy, O.A., Pavlou, P.A. and Venkatraman, N. (2013), "Digital business strategy: Toward a next generation of insights", MIS Quarterly, Vol. 37 No. 2, pp. 471-482.

Buhalis, D. and Amaranggana, A. (2014), "Smart tourism destinations", in Xiang, Z. and Tussyadiah, I. (Eds.), Information and Communication Technologies in Tourism 2014, Springer, Cham, pp. 553-564.

Carlucci, D., Lerro, A., Santarsiero, F., Lagrutta, R. and Orsi, V. (2025), "Assessing digital maturity in rural tourism: Insights from pilot areas in Basilicata", Proceedings IFKAD 2025: Knowledge Futures: AI, Technology, and the New Business Paradigm.

Carlucci, D., Lerro, A., Santarsiero, F., Lagrutta, R., Orsi, V. and Andrulli, G. (2025), "Models and tools to support innovation processes in cultural tourism ecosystems: Evidences and implications for Basilicata", Proceedings IFKAD 2025: Knowledge Futures: AI, Technology, and the New Business Paradigm.

Carlucci, D., Santarsiero, F., Lagrutta, R., Lerro, A. and Schiuma, G. (2024), "Achieving digital excellence in cultural tourism: The Digital Maturity Assessment Model", Proceedings IFKAD 2024: Translating Knowledge into Innovation Dynamics, pp. 1759-1770.

Denicolai, S., Cioccarelli, G. and Zucchella, A. (2010), "Resource-based local development and networked core-competencies for tourism excellence", Tourism Management, Vol. 31 No. 2, pp. 260-266.

Foss, N.J. and Saebi, T. (2017), "Fifteen years of research on business model innovation: How far have we come, and where should we go?", Journal of Management, Vol. 43 No. 1, pp. 200-227.

Gretzel, U., Sigala, M., Xiang, Z. and Koo, C. (2015), "Smart tourism: Foundations and developments", Electronic Markets, Vol. 25 No. 3, pp. 179-188.

Hjalager, A.M. (2010), "A review of innovation research in tourism", Tourism Management, Vol. 31 No. 1, pp. 1-12.

Jacobides, M.G., Cennamo, C. and Gawer, A. (2018), "Towards a theory of ecosystems", Strategic Management Journal, Vol. 39 No. 8, pp. 2255-2276.

Lerro, A., Santarsiero, F., Carlucci, D., Lagrutta, R. and Schiuma, G. (2024), "Participatory events fostering innovation dynamics in cultural tourism: Designing and testing methods and tools", Proceedings IFKAD 2024: Translating Knowledge into Innovation Dynamics, pp. 1681-1695.

Nambisan, S., Lyytinen, K., Majchrzak, A. and Song, M. (2017), "Digital innovation management: Reinventing innovation management research in a digital world", MIS Quarterly, Vol. 41 No. 1, pp. 223-238.

Nonaka, I. and Takeuchi, H. (1995), The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation, Oxford University Press, New York.

Santarsiero, F., Carlucci, D., Lerro, A. and Schiuma, G. (2024), "Navigating digital transformation and business model innovation in the tourism sector: Challenges, opportunities, and leadership styles", Measuring Business Excellence, Vol. 28 No. 3/4, pp. 426-438.

Teece, D.J. (2010), "Business models, business strategy and innovation", Long Range Planning, Vol. 43 No. 2/3, pp. 172-194.

Verhoef, P.C., Broekhuizen, T., Bart, Y., Bhattacharya, A., Dong, J.Q., Fabian, N. and Haenlein, M. (2021), "Digital transformation: A multidisciplinary reflection and research agenda", Journal of Business Research, Vol. 122, pp. 889-901.

Vial, G. (2019), "Understanding digital transformation: A review and a research agenda", The Journal of Strategic Information Systems, Vol. 28 No. 2, pp. 118-144.

Artificial Intelligence in Supply Chains: A Systematic Review and Knowledge Mapping of Research Trends

Benedetta Pini[1], Laura Monferdini[1], Barbara Bigliardi[1,2], Eleonora Bottani[1]
[1] University of Parma, Parma, Italy
[2] University of the Republic of San Marino, Dogana, San Marino

Abstract

Over the last decades, supply chain management has undergone a profound transformation, driven not only by rapid technological advancements but also by the increasing complexity and globalization of markets, the growing volatility of demand, and the emergence of new competitive, environmental, and institutional pressures that require greater efficiency, transparency, and adaptability. In this context, Artificial Intelligence (AI) has emerged as a key enabler of supply chain (SC) transformation, offering significant improvements in decision-making processes the optimization of the operations the enhancement of efficiency, resilience and sustainability. Despite the growing academic and managerial interest, the literature on AI in SCs remains fragmented across disciplines, sectors, and application domains.

This study aims to provide a systematic and integrated overview of the role of AI in SCs by identifying its main areas of application, emerging research trajectories, and key challenges associated with its adoption. To achieve the results, the research adopts a systematic literature review approach configured as a meta-review, focusing exclusively on review-type articles. The study adopts a systematic literature review approach configured as a meta-review, based on review articles retrieved from the Scopus database. The analysis integrates descriptive statistics with bibliometric techniques supported by VOSviewer.

The results reveal a non-uniform distribution of research, with a strong concentration in specific application domains, particularly agri-food, Supply chain-related areas such as logistics and distribution. AI applications in these domains primarily focus on demand forecasting, process optimization, traceability, and quality control. Emerging trends highlight the increasing relevance of sustainability, resilience, and digital integration, supported by complementary technologies such as blockchain, Internet of Things (IoT), and big data. At the same time, significant gaps persist, including the uneven maturity of AI applications across sectors, with some areas showing advanced implementation while others remain at an early stage. Limited technological integration, and insufficient attention to ethical and governance issues emerged.

Overall, the findings suggest that the main challenge is not the lack of interest in AI, but the need for strategies to fully exploit the potential of these technologies in managing various aspects of the SC. Based on this consideration, further research is needed to develop approaches capable of addressing the adoption of AI.

Keywords: Artificial Intelligence (AI), supply chain, Innovation, Industry 4.0 (I4.0), systematic literature review

Paper type: Academic Research Paper

1 Introduction

Over the last decades, supply chain management has undergone a profound transformation, driven by rapid technological advancements, the increasing complexity of global markets, and the emergence of new competitive, environmental, and institutional pressures (Christopher, 2022; Mentzer et al., 2001). In this evolving landscape, organizations are required to manage highly interconnected and dynamic systems, where efficiency alone is no longer sufficient. Instead, supply chains (SCs) must simultaneously ensure resilience, adaptability, transparency, and sustainability in response to growing uncertainty and disruption (Ivanov, 2019; Dubey et al., 2020).

In this context, Artificial Intelligence (AI) has progressively emerged as one of the most promising and disruptive technologies, capable of reshaping operational models, decision-making processes, and coordination mechanisms across the entire SC (Queiroz et al., 2020). AI-driven solutions enable firms to process large volumes of data, generate predictive insights, and support complex decisions in real time, thereby enhancing responsiveness and improving overall system performance (Baryannis et al., 2019). As a result, AI is increasingly recognized not merely as a technological tool, but as a strategic enabler of SC transformation.

The evolution of SCs is closely interlinked with the major stages of the industrial revolutions. The First Industrial Revolution laid the foundations for coordinating production and logistics flows through mechanization and the development of transportation infrastructures. Subsequently, the Second Industrial Revolution introduced mass production, electrification, and vertical integration, enabling SCs to expand beyond local boundaries and evolve into more structured systems (Chopra & Meindl, 2001). A major turning point occurred with the Third Industrial Revolution, characterized by the introduction of digitalization, automation, and

integrated information systems, which significantly improved efficiency, visibility, and coordination across SC processes (Efthymiou et al., 2021). Building on these developments, the Fourth Industrial Revolution, or Industry 4.0, has further accelerated this transformation by fostering interconnected, intelligent, and data-driven SCs through the integration of advanced technologies such as the Internet of Things (IoT), Big Data, blockchain, and AI (Frank et al., 2019; Ivanov et al., 2019).

In this scenario, AI plays a central role as an enabling technology for the creation, processing, and exploitation of knowledge along the SC (Sharma et al., 2024). Machine learning algorithms and predictive analytics techniques support complex decision-making processes, enhance planning capabilities, optimize inventory management and logistics flows, strengthen traceability, and improve responsiveness in uncertain and high-risk environments (Min, 2010; Baryannis et al., 2019).

More recently, the emergence of the Industry 5.0 paradigm has further expanded this perspective by placing human centricity, environmental sustainability, and system resilience at the core of industrial development. From this perspective, AI is no longer just a tool for automation, but also supportshuman–machine collaboration, circular economy practices and value creation across SCs (Longo et al., 2020; Nahavandi, 2019).

Despite growing academic and managerial interest, the adoption of AI in SCs is not without challenges. The literature highlights several critical barriers, including high implementation costs, shortages of specialized skills, difficulties in integrating AI solutions with legacy systems, and organizational resistance to change (Dwivedi et al., 2021; Kshetri, 2021). Furthermore, various ethical issues (e.g., data privacy, algorithmic bias, and transparency) may hinder the effective integration of AI into strategic processes.

Based on the above, this study aims to provide an integrated and critical overview of the state of the art on the application of AI in SCs through a systematic literature review approach. Specifically, the research is designed as a meta-review focusing exclusively on review-type articles, with the objective of synthesizing, comparing, and interpreting consolidated scientific contributions in the field (Tranfield et al., 2003).

The paper is structured as follows. Section 2 presents the objectives and the methodologies adopted in order to achieve the results, described in section 3. Finally, the conclusions along with the limitations and the future research of the study are presented.

2 Objective and Methodology

The present study aims to systematically analyze the role of AI in SCs, with the objective of understanding its main application domains, the emerging trajectories of research, and the implications arising from its adoption across different organizational and industrial contexts. More specifically, the study investigates how AI is applied along SC processes, which technological development paths are shaping the scientific debate, and how these applications relate to key issues such as sustainability, digital transformation, risk management, resilience, and ethical concerns. From this perspective, the paper seeks not only to reconstruct the current state of the literature, but also to identify the main knowledge gaps and the areas that require further theoretical and empirical investigation.

To address these objectives, the research is based on a literature analysis combining descriptive and bibliometric approaches, in order to map the structure of the field and identify the connections among its main emerging themes. This methodological choice is particularly suitable for investigating a rapidly growing and fragmented research domain, such as AI in SCs, where contributions are dispersed across different disciplines, sectors, and technological perspectives (Donthu et al., 2021; Zupic & Čater, 2015).

The literature review was designed as a *review of reviews (Grant and Booth, 2009)*, focusing exclusively on review-type articles. This choice responds to the need to provide a structured, critical, and comparative overview of an expanding body of knowledge by integrating already consolidated scientific contributions. In this way, the analysis goes beyond a mere summary of existing studies and aims instead to interpret and connect them, so as to highlight convergences, divergences, recurring thematic cores, and persistent gaps in the debate. Such an approach is particularly appropriate in a field like AI applied to SCs, which is characterized by a strong heterogeneity of perspectives, application contexts, and technologies (Tranfield et al., 2003).

The data collection process was carried out using the Scopus database, selected for its broad interdisciplinary coverage, the quality of its metadata, and the possibility of applying advanced filters to ensure rigor and transparency in the selection procedure. The corpus was built through an initial query based on the terms "*artificial intelligence*" and "*supply chain*", searched within titles, abstracts, and keywords.

This first search returned 4,419 documents. The dataset was then progressively refined according to inclusion criteria consistent with the objectives of the study. First, the search was restricted to review articles only, reducing the number of documents to 303. Second, only English-language publications were retained, which reduced the sample to 296 documents. Finally, works published in 2025 were excluded in order to avoid the

inclusion of overly recent contributions that may still be incomplete or unstable. The final corpus therefore consisted of 290 documents. The final search string was as follows:

TITLE-ABS-KEY ("artificial intelligence" AND "supply chain") AND (LIMIT-TO (DOCTYPE, "re")) AND (LIMIT-TO (LANGUAGE, "English")) AND (EXCLUDE (PUBYEAR, 2025))

Once the corpus was defined, the documents were analyzed by combining a descriptive examination of bibliographic metadata with a bibliometric analysis supported by the software VOSviewer. The descriptive analysis examine year of publication, most productive authors, top journals, and the distribution of contributions across different application domains. The bibliometric component, in turn, enabled a deeper investigation of the intellectual structure of the field through network analysis, which is particularly useful for mapping the main nodes of the literature and the relationships among them (van Eck & Waltman, 2010).

More specifically, the analysis focused on the co-occurrence of keywords, with the aim of identifying the main thematic clusters, the most central concepts, and the principal evolutionary trajectories of the scientific debate. In this sense, the integration of descriptive and bibliometric analysis allowed both a quantitative mapping of the field and a qualitative interpretation of its internal dynamics and conceptual development (Donthu et al., 2021; Zupic & Čater, 2015).

3 Results

3.1 Descriptive analysis

Figure 1 shows the trend over time in scientific publications focusing on the application of AI in SCs. The data reveal a significant increase in the number of from 2019 onward and a peak reached in 2024. This trend indicates the progressive consolidation of the topic within both academic and managerial debates, as well as the growing recognition of AI as a strategic lever for SC transformation.

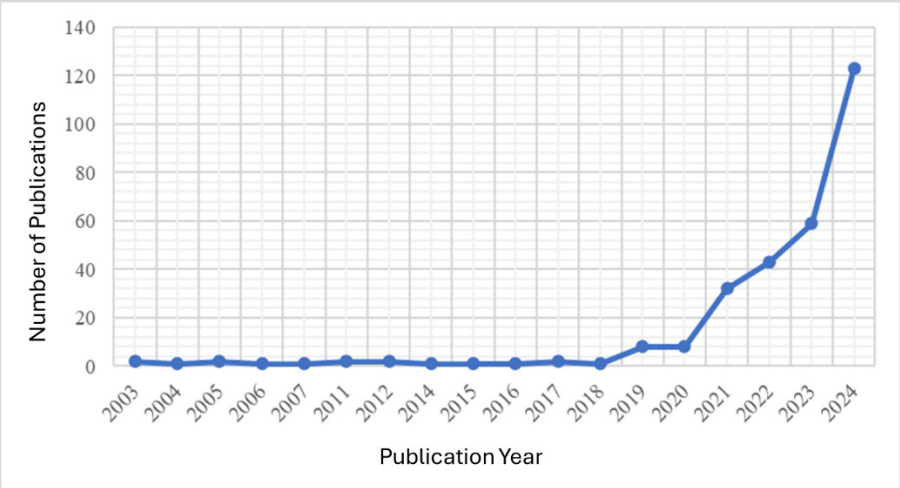


Figure 1. Number of publications by year.

The earliest contribution identified in the analyzed corpus is the article by Choy et al. (2003), which focuses on an intelligent system for supplier selection and benchmarking. This work can be considered pioneering, as it anticipated the use of AI as a support tool for decision-making processes in SC management, particularly in procurement activities. At a time when the adoption of intelligent technologies was still limited, this study had already introduced the potential of AI to improve transparency, efficiency, and objectivity in supplier evaluation decisions.

Until 2019, the annual number of publications remained relatively limited, suggesting an early stage in the development of the field, during which AI applications in SCs were still constrained by technological, infrastructural, and organizational limitations. From 2019 onward, however, the number of studies increased exponentially. This acceleration appears to be attributable to several factors. First, the rapid advancement of techniques such as machine learning, deep learning, and predictive analytics significantly expanded the application potential of AI in SC processes. Second, recent global crises, particularly the COVID-19 pandemic, highlighted the vulnerability of SCs and reinforced interest in intelligent solutions capable of improving

resilience, visibility, and predictive capacity. Finally, growing institutional and industrial support for digitalization fostered investments and experimentation in logistics and production settings.

Overall, the temporal dynamics of publications confirm that AI has evolved from an emerging topic into a consolidated research stream, becoming one of the principal drivers of digital transformation in SCs.

Figure 2 presents the main authors by number of publications, considering those with at least three contributions in the analyzed corpus.

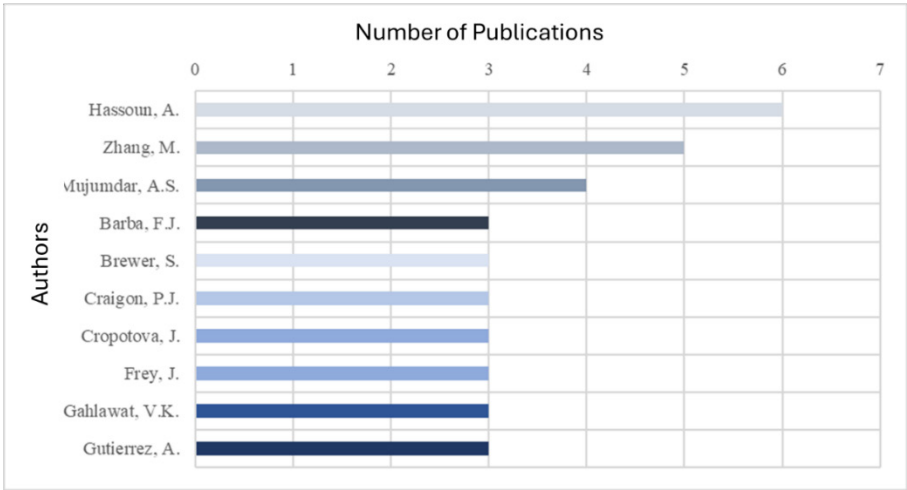


Figure 2. Top authors

The analysis reveals the presence of several particularly active scholars whose work reflects the main thematic orientations of the literature. In particular, a core group of contributions is strongly concentrated in the agri-food sector, confirming the central role of this sector in research on AI. These studies mainly explore the integration of AI, blockchain, the IoT, digital twins, and smart packaging for monitoring quality, safety, traceability, and logistics optimization in food SCs.

Alongside this stream, a second group of studies highlights the growing attention devoted to the ethical and governance implications of adopting intelligent technologies in SCs. These contributions emphasize issues such as transparency, accountability, data management, and the need for shared ethical languages among SC actors., showing that AI adoption is not only a technical matter, but also raises important social and ethical issues.

A further stream concerns the use of AI to address major challenges related to sustainability, resilience, and innovation in agri-food and manufacturing systems. In this context, digital technologies are increasingly regarded as essential tools for coping with geopolitical instability, climate change, health crises, and growing pressure toward more sustainable production models. Overall, the literature appears to be moving beyond isolated technological applications and is progressively framing AI as a cross-cutting driver of organizational transformation, sustainability, and systemic innovation.

Figure 3 shows the leading scientific journals by number of publications. The analysis highlights a strong heterogeneity of publication outlets, which reflects the interdisciplinary nature of the topic. Among the most relevant journals are *Sustainability*, *Trends in Food Science and Technology*, *International Journal of Production Research*, *Heliyon*, *Logistics*, *Computers and Electronics in Agriculture*, *Energies*, *Foods*, *IEEE Access*, and *Science of the Total Environment*. The distribution of publications across these outlets makes it possible to identify three main areas of convergence.

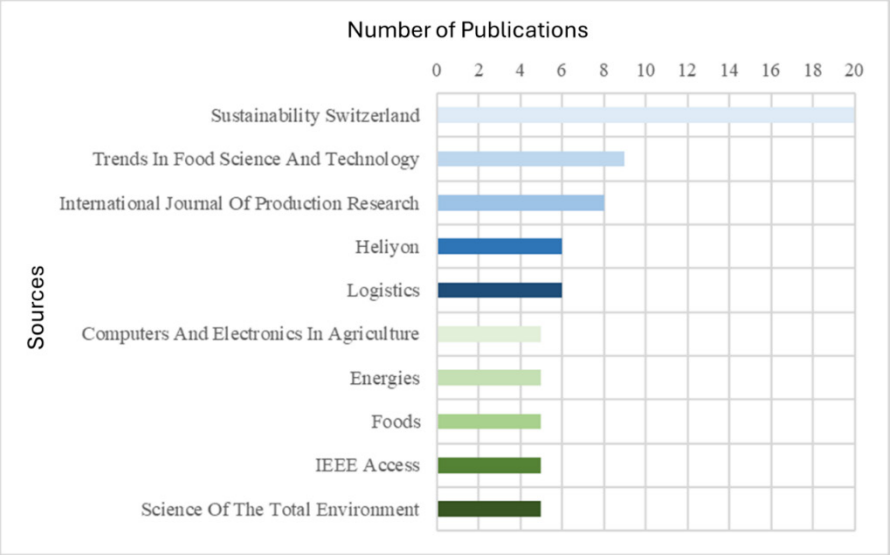


Figure 3. Top sources.

Sustainability concerns the link between technologies and environmental issues. The second (i.e., *Trends in Food Science and Technology*) is more closely related to the technological innovation in agri-food systems, where AI is studied in relation to quality, safety, traceability, and waste reduction. The third (i.e., *International Journal of Production Research*) deals with the more strictly managerial and engineering dimension, with studies focusing on process optimization, forecasting, risk management, and automation. Taken together, these journals demonstrate that research on AI in SCs lies at the intersection of management, technology, sustainability, and applied sciences.

Figure 4 presents the distribution of publications across application domains. The results show a clear predominance of the food and agriculture industry, which represents the sector with the highest number of contributions. In this domain, AI is mainly applied to quality and food safety monitoring, optimization of production and agricultural processes, waste reduction, automation in precision agriculture, and the improvement of product traceability along the SC.

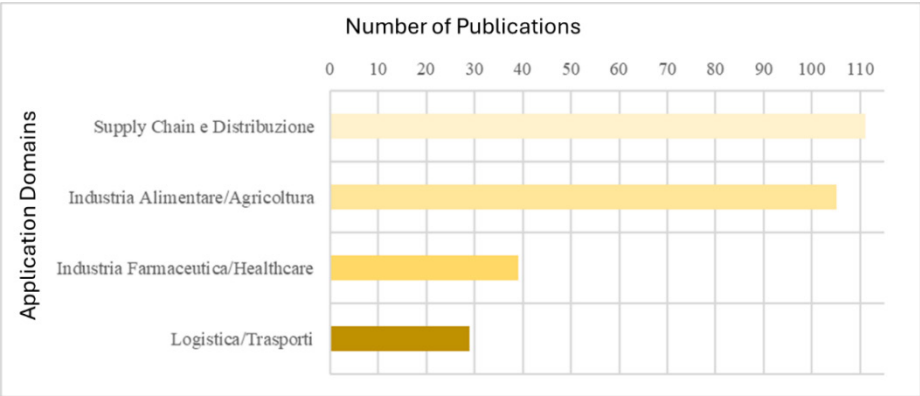


Figure 4. Main application domains.

A second particularly relevant domain concerns SC and distribution management in a broader sense. In these studies, AI is mainly employed for demand forecasting, inventory optimization, logistics flow planning, and the enhancement of transparency and resilience through integration with technologies such as blockchain and machine learning. This is followed by the pharmaceutical/healthcare and logistics/transportation areas, which also show a considerable number of contributions and growing academic interest. In these contexts, the main applications concern monitoring environmental conditions along the SC, forecasting the demand for drugs and medical devices, supporting distribution in emergency situations, and optimizing transport operations.

This analysis confirms that AI is now recognized as a strategic lever for improving efficiency, resilience, and sustainability across different industrial contexts. At the same time, it reveals a concentration of research in certain areas, whereas others remain less developed, thus suggesting opportunities for further investigation.

3.2 Network analysis

Figure 5 presents the co-occurrence map of author keywords generated through VOSviewer. In the middle of the network the terms “artificial intelligence”, “machine learning”, and “SC management” appear; these constitute the main nodes of the field’s thematic structure. This centrality is partly because these terms reflect the query used to construct the sample of papers under analysis; however, they also act as conceptual nodes capable of connecting different sub-streams of the literature.

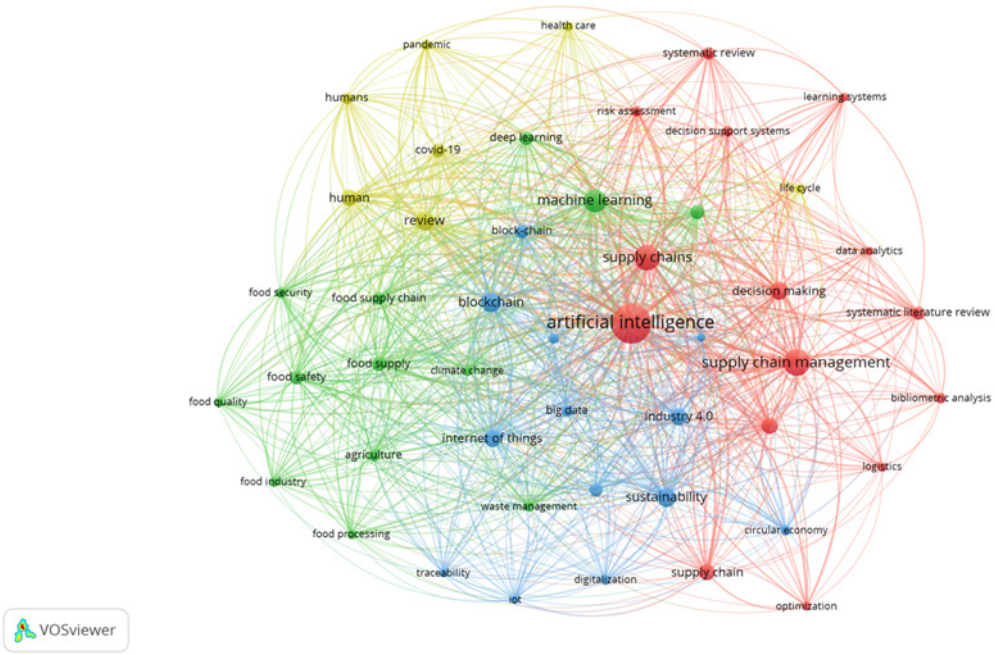


Figure 5. Co-occurence mapping of author keywords (source: VOSviewer).

Several thematic clusters are organized around these central nodes. A first cluster gathers enabling digital technologies such as IoT, blockchain, big data, and deep learning, highlighting the close relationship between AI and broader digitalization processes in SCs. From this perspective, AI does not emerge as an isolated technology, but rather as a component of an integrated digital ecosystem in which data availability, connectivity, and interoperability among systems play a fundamental role.

A second cluster concerns operational applications and the main improvement goals of SCs, with keywords such as traceability, food safety, quality control, and logistics. The proximity of these terms to the central nodes indicates that the literature mainly considers AI adoption as a lever for improving safety, transparency, reliability, and efficiency along the value chain, with particular relevance in the agri-food sector.

A third thematic core is formed by sustainability-related terms, including sustainability, environmental impact, and resource optimization. Their presence and interconnection with the technological and managerial clusters suggest that more recent research no longer interprets AI solely as a tool for improving efficiency, but increasingly as an enabling factor for sustainable, resilient, and resource-efficient SCs. In this sense, the map confirms the progressive integration of the digital agenda with the sustainability agenda.

The network also reveals a marked interdisciplinarity. Terms related to computer science, industrial engineering, logistics, food safety, and environmental studies are strongly intertwined, demonstrating that innovation in this field is rooted in the interaction among different theoretical perspectives and competencies. Keywords such as decision support systems, predictive analytics, and automation further reinforce the idea that one of AI’s main contributions to SCs lies in its ability to support decision-making, anticipate disruptions, and improve coordination across actors and activities.

Figure 6, which presents the temporal overlay visualization of keyword co-occurrences, makes it possible to reconstruct the evolution of the field over time. The terms displayed in blue-violet, corresponding to earlier streams, mainly concern the classical concepts of the discipline, such as SC management, logistics, optimization, decision support systems, and systematic review. This initial stage reflects a predominantly managerial orientation focused on operational efficiency and process optimization.

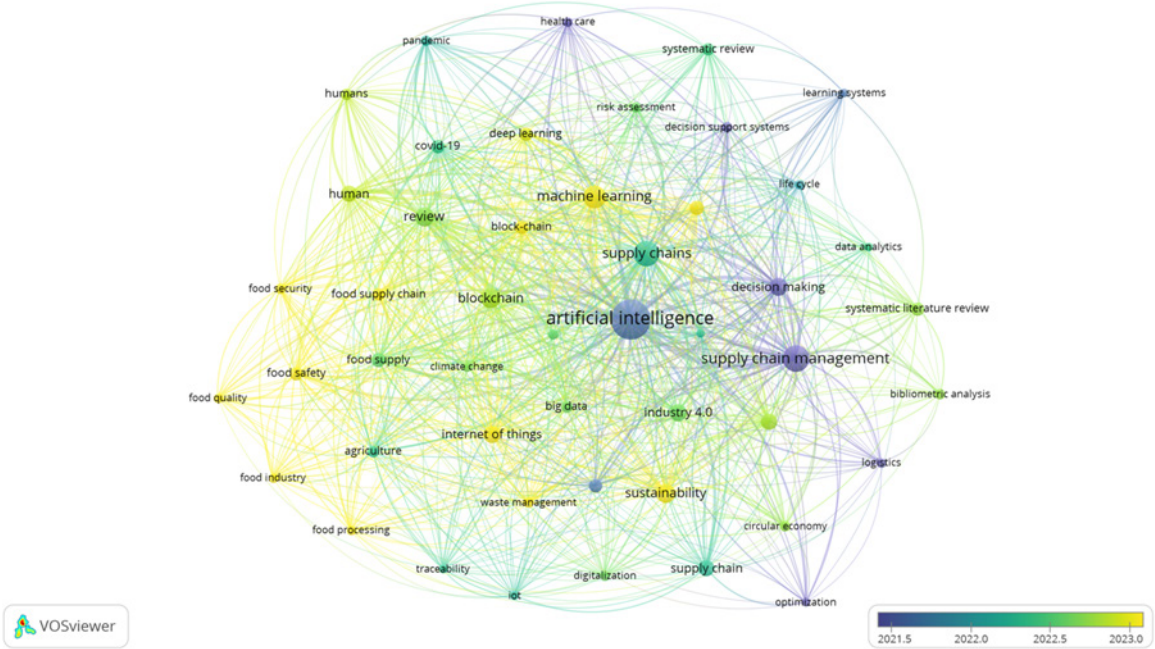


Figure 6. Network Analysis – Keywords Co-Occurrence Analysis Overlay Visualization (source: VOSviewer).

In the intermediate phase, between 2021 and 2022, terms such as artificial intelligence, machine learning, data analytics, industry 4.0, digitalization, risk assessment, and decision making become more prominent. This configuration signals a transition toward increasingly data-driven and predictive SCs, in which AI begins to be understood not merely as a tool for automation, but as a cognitive architecture supporting the analytical, learning, and adaptive capabilities of organizations.

In the most recent part of the map, highlighted in yellow, attention shifts toward themes such as sustainability, circular economy, waste management, and traceability, often associated with specific application domains, particularly agri-food. At the same time, complementary technologies such as blockchain, Internet of Things, and big data become more visible, as they are considered essential for ensuring data integrity, visibility, and reliability along the SC. The term associated with the pandemic context also remains present, having accelerated interest in resilience, digitalization, and adaptive capacity in SCs.

Overall, the temporal overlay reveals a clear evolutionary trajectory: from an initial phase centered on optimization and decision support, the field has moved toward a progressive integration of AI, machine learning, and analytics, eventually reaching a growing emphasis on sustainability, transparency, traceability, and advanced sectoral applications. AI remains the conceptual core of the field, but the most dynamic and recent themes are those that connect digitalization, sustainability, and SC governance.

4 Conclusion, Limits and Future research Agenda

This study provides a systematic and integrated overview of the role of AI in SCs through a meta-review approach supported by bibliometric techniques. The findings highlight how AI is increasingly shaping the transformation of SCs by enhancing operational efficiency, predictive capabilities, resilience, and transparency. In particular, the analysis reveals a strong concentration of research in specific application domains, such as the agri-food sector, SC and distribution management, and, more recently, healthcare and logistics.

Overall, the results confirm that AI should no longer be interpreted merely as a technological tool, but rather as a strategic enabler capable of supporting the transition toward data-driven, interconnected, and value-oriented SC models. At the same time, the literature shows a gradual shift from efficiency-oriented perspectives toward broader approaches that increasingly incorporate sustainability, traceability, and risk management. However, such evolution remains uneven and fragmented, indicating the need for further theoretical and empirical development.

From a theoretical perspective, this study contributes by systematizing a fragmented body of knowledge, identifying key thematic clusters, and reconstructing the intellectual structure and evolution of the field. From a managerial standpoint, the findings suggest that AI adoption requires not only technological investments but also organizational transformation, development of new competencies, and stronger integration across systems

and actors within the SC. A systemic perspective is therefore essential to jointly address efficiency, sustainability, and resilience.

Despite its contributions, this study presents some limitations. First, the analysis is based exclusively on review articles. While this choice ensures a high level of abstraction and allows for the synthesis of consolidated knowledge, it may limit the inclusion of recent empirical insights emerging from primary studies. Second, the use of a single database (Scopus), although comprehensive and widely recognized, may have excluded relevant contributions indexed in other databases. Thirdly, whilst the bibliometric approach is effective in identifying research patterns, groups and trajectories, it fails to fully capture the qualitative and contextual aspects of AI implementation in real-world data science contexts. Consequently, the findings should be interpreted as a structured overview rather than an exhaustive representation of all existing evidence.

The results of this study open several promising future research. First, although sustainability is increasingly present in the literature, many studies still prioritize operational efficiency and cost reduction. Future research should more deeply explore the potential of AI to address environmental challenges in a systemic way, including emission reduction, circular economy models, and resource optimization across the entire SC.

Second, the findings reveal a persistent gap in technological integration. While AI is often discussed alongside technologies such as blockchain, IoT, and big data, their actual integration in practice remains fragmented. Future studies should investigate issues related to interoperability, standardization, and scalability, which are critical for enabling fully interconnected and intelligent SCs.

Third, there is a need for more interdisciplinary research approaches. Much of the existing literature focuses on isolated aspects of SC management, whereas future research should adopt holistic perspectives that consider the interaction between technological, organizational, social, and environmental dimensions.

Another important direction concerns data availability and quality. Since AI applications are based on large volumes of reliable and structured data, future research should address challenges related to data governance, data sharing across different sectors.

Finally, greater attention should be devoted to the ethical and regulatory implications of AI adoption. Issues such as algorithmic transparency, accountability, data privacy, and the impact on employment remain underexplored but are crucial for fostering trust and ensuring responsible AI implementation.

In conclusion, the main limitation of the current research landscape does not lie in the lack of interest in AI, but rather in its uneven and fragmented development across sectors and perspectives. Future research should therefore move toward more integrated, interdisciplinary, and responsible approaches capable of aligning technological innovation with sustainability and long-term value creation.

5 Ethical Declaration

Ethical approval was not required for this study, as the research is based exclusively on secondary data and does not involve human participants, personal data, or any form of sensitive information.

6 AI Declaration

No artificial intelligence (AI) tools were used in the preparation of this manuscript.

References

Baryannis, G., Dani, S., & Antoniou, G. (2019). Predictive analytics and artificial intelligence in supply chain management: Review and implications for the future. Computers & Industrial Engineering, 137(106024).

Choi, T. M., Wallace, S. W., & Wang, Y. (2018). Big data analytics in operations management. *Production and operations management*, *27*(10), 1868-1883.

Chopra, S., & Meindl, P. (2001). Strategy, planning, and operation. Supply Chain Management, 15(5), 71-85.

Choy, K. L., Lee, W. B., & Lo, V. (2003). An intelligent supplier relationship management system for selecting and benchmarking suppliers. International Journal of Technology Management, 26(7), 717-742.

Christopher, M. (2022). Logistics and supply chain management. Pearson Uk.

Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., & Lim, W. M. (2021). How to conduct a bibliometric analysis: An overview and guidelines. Journal of business research, 133, 285-296.

Dubey, R., Gunasekaran, A., Childe, S. J., Bryde, D. J., Giannakis, M., Foropon, C., ... & Hazen, B. T. (2020). Big data analytics and artificial intelligence pathway to operational performance under the effects of entrepreneurial orientation and environmental dynamism: A study of manufacturing organisations. International journal of production economics, 226, 107599.

Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., ... & Williams, M. D. (2021). Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International journal of information management*, *57*, 101994.

Efthymiou, O. K., & Ponis, S. T. (2021). Industry 4.0 Technologies and Their Impact in Contemporary Logistics: A Systematic Literature Review. Sustainability, 13(21), 11643. https://doi.org/10.3390/su132111643

Frank, A. G., Dalenogare, L. S., & Ayala, N. F. (2019). Industry 4.0 technologies: Implementation patterns in manufacturing companies. International journal of production economics, 210, 15-26.

Grant, M.J. and Booth, A. (2009), A typology of reviews: an analysis of 14 review types and associated methodologies. Health Information & Libraries Journal, 26: 91-108. https://doi.org/10.1111/j.1471-1842.2009.00848.x

Ivanov, D. (2021). Supply chain viability and the COVID-19 pandemic: a conceptual and formal generalisation of four major adaptation strategies. International Journal of Production Research, 59(12), 3535-3552.

Ivanov, D., Dolgui, A., & Sokolov, B. (2019). The impact of digital technology and Industry 4.0 on the ripple effect and supply chain risk analytics. International journal of production research, 57(3), 829-846.

Ketchen Jr, D. J., & Hult, G. T. M. (2007). Bridging organization theory and supply chain management: The case of best value supply chains. *Journal of operations management*, *25*(2), 573-580.

Kshetri, N. (2021). Blockchain and sustainable supply chain management in developing countries. *International journal of information management*, *60*, 102376.

Longo, F., Padovano, A., & Umbrello, S. (2020). Value-oriented and ethical technology engineering in industry 5.0: A human-centric perspective for the design of the factory of the future. *Applied sciences*, *10*(12), 4182.

McElwee, M. (1998) "Chaos Theory and Complexity as Fountainheads for Design of an Organization Theory Building Workshop", Paper read at XIVth World Congress of the International Sociological Association, Montreal, Canada, July.

Mentzer, J. T., DeWitt, W., Keebler, J. S., Min, S., Nix, N. W., Smith, C. D., & Zacharia, Z. G. (2001). Defining supply chain management. Journal of Business logistics, 22(2), 1-25.

Min, H. (2010). Artificial intelligence in supply chain management: theory and applications. International Journal of Logistics: Research and Applications, 13(1), 13-39.

Nahavandi, S. (2019). Industry 5.0—A human-centric solution. *Sustainability*, *11*(16), 4371.

Qian, C., Murphy, S. I., Orsi, R. H., & Wiedmann, M. (2023). How can AI help improve food safety?. *Annual Review of Food Science and Technology*, *14*(1), 517-538.

Qian, J., Dai, B., Wang, B., Zha, Y., & Song, Q. (2022). Traceability in food processing: problems, methods, and performance evaluations—a review. *Critical Reviews in Food Science and Nutrition*, *62*(3), 679-692.

Queiroz, M. M., Telles, R., & Bonilla, S. H. (2020). Blockchain and supply chain management integration: a systematic review of the literature. Supply chain management: An international journal, 25(2), 241-254.

Sharma, P., Gunasekaran, A., & Subramanian, G. (2024). Enhancing supply chain: exploring and exploiting AI capabilities. *Journal of Computer Information Systems*, 1-15.

Tranfield, D., Denyer, D., & Smart, P. (2003). Towards a methodology for developing evidence-informed management knowledge by means of systematic review. *British journal of management*, *14*(3), 207-222.

Van Eck, N., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *scientometrics*, *84*(2), 523-538.

Zupic, I., & Čater, T. (2015). Bibliometric methods in management and organization. Organizational research methods, 18(3), 429-472.

When Lead-User Signals Mislead: Entrepreneurial Sensemaking in Sport-Tech Innovation

Barbara Bigliardi[1,2], Virginia Dolci[2], Serena Filippelli[1], Benedetta Pini[1]
[1] University of Parma, Parma, Italy
[2] University of San Marino, San Marino

Abstract

This study investigates how sport-tech entrepreneurs make sense of heterogeneous lead-user signals in new product development under conditions of uncertainty. While prior research has widely emphasized the role of lead users as sources of valuable and actionable knowledge, it often assumes that user inputs reduce uncertainty by providing clear guidance for innovation. Challenging this assumption, this study draws on sensemaking theory, signaling theory, and paradox theory to conceptualize lead-user feedback as a set of ambiguous and potentially conflicting signals that require interpretation rather than direct translation into product decisions.

An exploratory sequential research design was adopted. In the first phase, survey data from 97 elite and professional athletes were analyzed using k-means cluster analysis to identify distinct configurations of lead-user signals. The analysis revealed two main profiles: (i) cautious/ambivalent users, characterized by moderate perceived usefulness combined with concerns related to usability, comfort, adaptability, and willingness to engage in co-creation; and (ii) receptive/engaged users, characterized by consistently positive evaluations across all dimensions, including performance value, usability, data reliability, and collaborative intent. These findings highlight that lead-user input is not homogeneous, but instead structured in differentiated and sometimes contradictory configurations.

In the second phase, semi-structured interviews with three sport-tech entrepreneurs were conducted to explore how these configurations are interpreted in practice. The qualitative findings show that ambivalent signals are subject to divergent interpretations: they may be framed either as remediable design issues requiring iterative refinement or as indications of deeper misalignment between the product and its target users. Conversely, positive signals are generally interpreted as validation of the product concept, although they do not eliminate uncertainty and are not always associated with immediate readiness for scaling. Across cases, decision-making emerges as an iterative and interpretive process shaped by entrepreneurs' assumptions, strategic priorities, and technical constraints.

By integrating quantitative and qualitative evidence, this study contributes to lead-user research by reframing user inputs as signals that generate both insight and ambiguity. It also advances research on entrepreneurial decision-making by highlighting the central role of sensemaking in constructing coherence among heterogeneous evaluations. From a managerial perspective, the findings suggest that the key challenge is not merely to collect user feedback, but to develop capabilities to interpret and prioritize conflicting signals in technology-driven innovation contexts.

Keywords: Entrepreneurial sensemaking; Lead-user signals; Technological uncertainty; Entrepreneurial decision making; Sport-tech innovation

Paper type: Academic Research Paper

1 Introduction

New digital technologies are transforming the way individuals communicate, coordinate, and interact, increasingly relying on data-rich digital infrastructures embedded in products, services, and organizational processes (Porter and Heppelmann, 2015; Cusumano et al., 2021). In the sports domain, this transformation is particularly evident in the growth of sport-tech entrepreneurship, where technology-enabled solutions are used to support training, performance monitoring, athlete engagement, and competitive advantage (Ratten et al., 2021; Ratten, 2024). Sport-tech entrepreneurs are therefore expected not only to develop innovative digital products, but also to make product development decisions in contexts characterized by technological uncertainty, rapidly evolving user needs, and complex data-driven value propositions. These conditions make sport-tech innovation a particularly challenging decision-making context. Entrepreneurs often operate under volatile, uncertain, complex, and ambiguous conditions, where action must be taken with bounded rationality, limited resources, and time pressure, rather than through fully linear optimization of clearly defined alternatives (Kotlar and Sieger, 2019; Troise et al., 2022). In sport-tech manufacturing, these challenges are amplified because product innovation often unfolds through iterative and path-dependent processes involving problem framing, prototyping, field testing, and redesign. As a result, product development decisions rarely consist of discrete choices among stable alternatives; instead, they involve successive and interdependent decisions shaped by evolving technical constraints, user feedback, and market expectations.

To reduce uncertainty in early-stage product innovation, entrepreneurs often rely on lead users. Lead users are advanced users who experience needs ahead of the mainstream market and may therefore provide valuable insights into emerging problems, product requirements, and potential solutions (von Hippel, 1986). Prior research has emphasized that involving lead users can improve the novelty and market relevance of innovation outcomes, particularly under conditions of technological uncertainty (Yin et al., 2023). However, this perspective tends to assume that lead-user inputs provide relatively clear and actionable knowledge. Less attention has been paid to situations in which lead-user feedback generates ambiguity rather than clarity, especially when entrepreneurs must translate heterogeneous user inputs into manufacturable, usable, and scalable product architectures.

This issue is especially relevant in sport-tech innovation, where user feedback may simultaneously concern performance value, usability, comfort, data reliability, adaptability, and willingness to engage in further co-creation. These signals may not always align. For instance, athletes may perceive a technology as potentially useful while also questioning its comfort, integration into training routines, adaptability to specific needs, or reliability of data outputs. In such cases, lead-user feedback does not simply reduce uncertainty; rather, it may intensify interpretive demands by exposing entrepreneurs to multiple and potentially contradictory cues.

Grounded in sensemaking theory (Weick, 1995; Maitlis, 2005), this study conceptualizes entrepreneurial decision-making in sport-tech innovation as an interpretive process through which entrepreneurs notice, select, label, and act upon heterogeneous user signals. In addition, drawing on paradox theory (Smith and Lewis, 2011), we argue that such signals may activate persistent tensions in sport-tech product development, including performance enhancement versus comfort and intrusiveness, data richness versus trust and ownership, and elite customization versus scalability. Thus, rather than treating lead-user feedback as a direct input into product decisions, this study examines how entrepreneurs make sense of differentiated lead-user signals under uncertainty.

Accordingly, the study addresses the following research question:

How do sport-tech entrepreneurs make sense of heterogeneous lead-user signals and translate them into product development decisions under uncertainty?

Empirically, the study adopts an exploratory sequential design. In the first phase, survey data from 97 elite and professional athletes are used to identify distinct configurations of lead-user signals related to perceived performance value, adoption frictions, trust in data, and collaborative intent. In the second phase, an explanatory qualitative case study of the sport-tech venture involved in the development of the evaluated technology is used to explore how these differentiated signals are interpreted in practice and how they inform product development decisions.

2 Theoretical Background

2.1 Lead users and the limits of informational clarity

Lead users are widely recognized as a valuable source of innovation-relevant knowledge, as they experience needs ahead of the mainstream market and often develop solutions that anticipate broader demand (von Hippel, 1986; Schmid et al., 2022). Engaging with such users has been shown to enhance the novelty and market relevance of new product concepts, particularly under conditions of technological uncertainty and rapid change (Yin et al., 2023). Accordingly, lead users are typically conceptualized as providers of high-quality insights that firms can leverage to guide innovation processes (Schmid et al., 2022).

This perspective, however, implicitly assumes that inputs from lead users provide relatively clear and actionable guidance. User contributions are often treated as informative cues that can be directly translated into product features or design improvements (Poblete et al., 2023). Such an assumption becomes problematic in contexts characterized by high levels of uncertainty, complexity, and time pressure, where decision-makers operate under bounded rationality (Simon, 2000; Troise et al., 2022).

In these environments, increasing the amount of available information does not necessarily reduce uncertainty. Instead, user-generated inputs may be heterogeneous, reflecting different and sometimes conflicting evaluations of the same technology (Henkens et al., 2021). As a result, lead-user inputs may contribute to ambiguity rather than clarity, reinforcing the complexity of the informational environment (Eppler and Mengis, 2004).

Despite the growing emphasis on user involvement in innovation, limited attention has been paid to the conditions under which lead-user inputs generate ambiguity rather than actionable knowledge (Venesz et al., 2022). In particular, we still lack a clear understanding of how heterogeneous and potentially contradictory user signals affect decision-making in product development contexts, especially in technology-intensive domains

such as sport-tech, where multiple product attributes must be balanced simultaneously (Porter and Heppelmann, 2015; Raff et al., 2020).

This gap calls for a shift from viewing lead users as direct sources of innovation inputs to conceptualizing their contributions as signals whose interpretation is inherently uncertain. This perspective is particularly relevant in sport-tech innovation, where technologies must simultaneously deliver performance improvements, ensure usability in physically demanding contexts, and generate reliable data, thereby creating multiple and interdependent evaluation criteria (Henkens et al., 2021).

2.2 Lead-user signals and key dimensions of evaluation

In technology-intensive contexts such as sport-tech, entrepreneurs operate under conditions of uncertainty in which decision-making cannot be reduced to linear optimization processes (Kotlar & Sieger, 2019; Troise et al., 2022). Within such environments, inputs from users are more appropriately understood as signals that capture multiple and potentially conflicting evaluations of emerging technologies, rather than as straightforward information (Connelly et al., 2011).

Building on this perspective, we conceptualize lead-user inputs along four analytical dimensions: (i) perceived performance value, (ii) adoption frictions, (iii) trust in data, and (iv) collaborative intent, which capture distinct categories of signals relevant to product development decisions. Rather than aiming at exhaustiveness, this set reflects a minimal and theoretically grounded representation of the key sources of uncertainty that entrepreneurs must address when translating user inputs into product innovation choices.

These dimensions correspond to different, non-reducible decision problems and capture distinct categories of uncertainty related to desirability, feasibility, reliability, and user involvement. Perceived performance value concerns whether a technological solution generates meaningful performance improvements and thus justifies its development (Benar et al., 2024). Adoption frictions capture constraints related to usability, comfort, and integration within training routines, shaping the feasibility of use in practice (Kim et al., 2019). Trust in data refers to the perceived reliability and credibility of technological outputs, a critical issue in data-intensive environments where decision-making depends on the validity of information (Chamorro-Koc et al., 2021). Collaborative intent captures users' willingness to engage in co-creation processes, shaping the potential for iterative learning and product refinement (Prahalad & Ramaswamy, 2004; Zhang et al., 2020).

While analytically distinct, these dimensions are empirically interdependent and do not align in a linear manner. User evaluations may combine in heterogeneous and sometimes contradictory ways, generating configurations in which high perceived value coexists with significant frictions, or willingness to collaborate emerges alongside low trust in data. Such configurations are particularly salient in sport-tech innovation, where product development involves managing persistent tensions between performance and intrusiveness, data richness and reliability, and customization and scalability (Smith & Lewis, 2011; Henkens et al., 2021).

2.3 Entrepreneurial sensemaking of lead-user signals

Lead-user signals do not translate directly into product innovation decisions. Under conditions of uncertainty, decision-making relies on interpretive processes through which entrepreneurs assign meaning to heterogeneous and potentially conflicting cues (Weick, 1995; Maitlis, 2005).

In sport-tech contexts, such signals are not self-evident but require selection and interpretation within existing cognitive frames. This becomes particularly salient when signals combine in non-linear and contradictory ways, generating configurations that do not provide clear directions for action. In these situations, ambiguity stems not from a lack of information, but from the coexistence of multiple plausible interpretations.

Entrepreneurial decision-making thus involves constructing coherence among divergent evaluations in order to define what is feasible, valuable, and scalable within a given product trajectory (Iyer et al., 2022). This perspective shifts attention from the content of user inputs to the processes through which such inputs are interpreted and enacted in decision-making. Accordingly, the study adopts a two-stage design in which configurations of lead-user signals are first identified and then examined through qualitative insights from entrepreneurs.

3 Methodology

3.1 First phase: quantitative cluster analysis

This study adopts an exploratory sequential research design to investigate how lead-user signals inform, complicate, and potentially mislead entrepreneurial sensemaking in sport-tech innovation. This design is

appropriate because the study aims first to identify heterogeneous configurations of user-generated signals and then to interpret how such signals may be used by entrepreneurs when making product development decisions under uncertainty. In line with the theoretical framing of the study, entrepreneurial decision making is not treated as a purely linear evaluation of market information, but as a sensemaking process through which decision makers interpret ambiguous and sometimes contradictory cues emerging from advanced users. The empirical setting is the sport-tech domain, where product innovation is strongly shaped by digital technologies, data-rich infrastructures, and continuous interaction with expert users.

Consistent with the theoretical framework, the survey operationalized the four analytical dimensions of lead-user signals through six empirical variables as shown in Table 1.

Table 1. Operationalization of lead-user signal dimensions.

Theoretical dimension	Empirical variabile	Reference
Perceived performance value	Perceived usefulness	Bardus et al., 2021
Adoption frictions	Usability; Comfort and ergonomics; Adaptability	Brooke, 1996
Trust in data	Data quality and reliability	Bardus et al., 2021
Collaborative intent	Willingness to engage in future co-creation	Ramaswamy and Ozcan, 2018

Perceived performance value was captured through perceived usefulness; adoption frictions were captured through usability, comfort and ergonomics, and adaptability; trust in data was captured through data quality and reliability; and collaborative intent was captured through willingness to engage in future co-creation. The use of multiple indicators for adoption frictions reflects the specific nature of sport-tech innovation, where the practical integration of technologies depends not only on ease of use, but also on bodily comfort, ergonomic fit, and adaptability to heterogeneous training routines and athletic needs.

The first phase of the study was based on a quantitative survey administered to 97 elite and professional athletes. The survey captured athletes' evaluations of a sport-tech innovation across six dimensions: perceived usefulness, usability, comfort and ergonomics, data quality and reliability, adaptability, and willingness to engage in future co-creation. These variables were selected to reflect different types of lead-user signals:

- perceived usefulness captures the expected performance value of the technology (Bardus et al., 2021);
- usability and comfort/ergonomics capture potential adoption frictions (Brooke, 1996);
- data quality and reliability capture trust-related concerns (Bardus et al., 2021);
- adaptability reflects the extent to which the technology can fit heterogeneous athletic needs (Brooke, 1996);
- and willingness to engage in future co-creation captures athletes' collaborative intent (Ramaswamy and Ozcan, 2018).

To identify distinct configurations of lead-user signals, a k-means cluster analysis was performed using SPSS software. The analysis was conducted on the six survey variables listed above, with a two-cluster solution specified. Cases were handled using listwise deletion, although no missing cases were reported in the final analysis. The use of k-means clustering is consistent with the exploratory nature of the study. Rather than testing causal relationships among variables, the aim was to classify athletes into meaningful profiles based on the pattern of their evaluations. This approach allows the study to move beyond average-level interpretations of user feedback and to examine whether lead users communicate homogeneous or differentiated signals to entrepreneurs. In this respect, the cluster analysis serves as a bridge between the quantitative and qualitative phases of the research: it identifies empirical signal configurations that can later be used to interpret how entrepreneurs make sense of user feedback during product development. The second phase of the study is qualitative and complementary. Semi-structured interviews with sport-tech entrepreneurs are used to explore how the differentiated signals emerging from the athlete survey are interpreted in practice. This phase focuses on the cognitive and interpretive processes through which entrepreneurs notice, select, label, and act upon ambiguous user inputs.

3.2 Second phase: qualitative analysis

The second phase is qualitative and complementary, and examines how entrepreneurs interpret the differentiated configurations of lead-user signals identified in the quantitative analysis. Consistent with the exploratory sequential design, the aim is not to generalize findings, but to provide process-level insights into how such signals are interpreted and translated into product development decisions under uncertainty.

Data were collected through semi-structured interviews with three sport-tech entrepreneurs directly involved in product development and commercialization (Yin, 2009). Informants were purposively selected to

ensure alignment between the user-generated signals captured in the first phase and the interpretive processes examined in the second phase. Given the exploratory nature of this phase, the goal was not theoretical saturation, but to capture variation in how similar signal configurations are interpreted.

The interview protocol focused on how entrepreneurs interpret and act upon heterogeneous user feedback, with particular attention to situations characterized by ambiguity or conflicting evaluations. Questions addressed how feedback is collected and prioritized, how different signal dimensions are interpreted, how tensions are managed, and how these interpretations inform product decisions.

Data were analyzed using an abductive thematic approach. Initial coding was guided by the four dimensions of lead-user signals, while remaining open to emergent patterns (Yin, 2009). A second level of analysis focused on sensemaking processes, examining how signals were interpreted as indicators of product validation, design limitations, adoption barriers, or ambiguity, and how interpretations varied across respondents.

This approach enables integration between the two phases: while cluster analysis identifies configurations of lead-user signals, the qualitative inquiry explains how such configurations are interpreted in practice, highlighting the role of sensemaking in entrepreneurial decision-making.

4 Results

4.1 Cluster analysis

To explore whether elite and professional athletes convey homogeneous or differentiated signals regarding sport-tech innovation, a k-means cluster analysis was conducted on six dimensions capturing athletes' evaluations of the product concept: perceived usefulness, usability, comfort and ergonomics, data quality and reliability, adaptability, and willingness to engage in future co-creation.

The two-cluster solution identified two distinct profiles of lead-user signals (Table 2). The first cluster included 33 athletes, corresponding to approximately 34% of the sample, whereas the second cluster included 64 athletes, corresponding to approximately 66% of the sample.

The first cluster can be interpreted as a group of cautious or ambivalent lead users. Athletes in this cluster reported moderate perceived usefulness of the technology, but lower evaluations on the more implementation-related dimensions. Specifically, the cluster center for perceived usefulness was 3.3, while usability was 3.0, comfort and ergonomics 2.8, data quality and reliability 3.2, adaptability 2.3, and willingness to engage in future co-creation 2.85. This profile suggests that these athletes do not reject the technology outright; rather, they recognize some potential value while simultaneously expressing reservations about its practical integration into training routines, its adaptability to specific needs, and their willingness to remain involved in future co-development activities. In this sense, their signals are not simply negative, but ambiguous: they combine moderate value recognition with concerns related to usability, comfort, adaptability, and continued collaboration.

The second cluster can instead be interpreted as a group of receptive and engaged lead users. Athletes in this cluster reported higher scores across all six dimensions. The cluster centers were 4.1 for perceived usefulness, 3.5 for usability, 3.8 for comfort and ergonomics, 3.9 for data quality and reliability, 3.4 for adaptability, and 3.60 for willingness to engage in future co-creation. This profile indicates a more favorable configuration of lead-user signals, in which the perceived performance value of the technology is accompanied by comparatively stronger evaluations of its usability, comfort, data reliability, and adaptability. Importantly, this group also showed greater willingness to participate in future co-creation, suggesting that these athletes may represent a more immediately actionable source of user knowledge for entrepreneurial product development.

Table 2. Cluster centers and profiles of lead-user signals

Variable	Cluster 1: Cautious / ambivalent lead users	Cluster 2: Receptive / engaged lead users
Perceived usefulness	3.3	4.1
Usability	3.0	3.5
Comfort and ergonomics	2.8	3.8
Data quality and reliability	3.2	3.9
Adaptability	2.3	3.4
Willingness to engage in future co-creation	2.85	3.60
N	33	64
% of sample	34.0%	66.0%

The descriptive ANOVA results show that the two clusters differ across all six dimensions. The largest difference emerged for adaptability, $F(1,95) = 64.585$, $p < .001$, followed by perceived usefulness, $F(1,95) = 42.038$, $p < .001$, comfort and ergonomics, $F(1,95) = 39.237$, $p < .001$, and data quality and reliability, $F(1,95) = 36.517$, $p < .001$. Differences were also observed for usability, $F(1,95) = 14.152$, $p < .001$, and willingness to engage in future co-creation, $F(1,95) = 10.610$, $p = .002$. However, in line with the SPSS output, these F-tests should be interpreted only descriptively, since the clusters were created to maximize between-group differences and the observed significance levels are therefore not suitable for formal hypothesis testing. Overall, the results suggest that lead-user input in sport-tech innovation is not uniform. While a majority of athletes provided broadly favorable signals, a substantial minority expressed a more ambivalent pattern, combining moderate perceived usefulness with lower evaluations of adaptability, comfort, usability, and future co-creation willingness. This finding is theoretically relevant because it qualifies the assumption that lead users necessarily reduce uncertainty for entrepreneurs. Instead, the analysis shows that lead users may generate different and partly conflicting signals: some signals point to product potential, while others reveal frictions associated with use, integration, and scalability. These configurations therefore create interpretive demands for entrepreneurs, who must decide whether cautious feedback reflects temporary product limitations, deeper adoption barriers, or valuable indications for redesign.

From a sensemaking perspective, these results indicate that entrepreneurial decision makers are not simply receiving clear market information from advanced users. Rather, they are exposed to heterogeneous cues that require interpretation, selection, and prioritization. The coexistence of receptive and ambivalent lead-user profiles may intensify the paradoxes already embedded in sport-tech manufacturing, such as the tension between performance enhancement and comfort, customization and scalability, and data richness and user trust. Therefore, the cluster analysis provides an empirical basis for the subsequent qualitative phase, where interviews with sport-tech entrepreneurs can explain how such differentiated lead-user signals are noticed, interpreted, and translated into product development decisions.

4.2 Qualitative insights: interpreting lead-user signal configurations

To complement the cluster analysis, qualitative insights were collected to explore how sport-tech entrepreneurs interpret differentiated configurations of lead-user signals in practice. The interviews focused on how decision makers make sense of heterogeneous and potentially conflicting user feedback, and how such interpretations inform product development decisions.

Consistent with the quantitative findings, entrepreneurs do not interpret lead-user signals in a uniform way. In relation to the ambivalent profile identified in Cluster 1, respondents offered partially divergent interpretations. Some entrepreneurs tended to frame moderate evaluations combined with usability and adaptability concerns as temporary and remediable issues. As noted by one informant, such feedback is often interpreted as “problems to be solved” rather than as fundamental limitations of the product. This perspective reflects an underlying assumption that perceived value justifies further refinement efforts.

However, not all respondents shared this interpretation. Another entrepreneur described similar feedback as potentially indicative of a misalignment between the product and its intended users, suggesting that such signals may reflect “non-target users” rather than correctable design issues. This divergence highlights how the same configuration of signals can support different interpretations, ranging from opportunities for incremental improvement to indications of more structural market misfit.

With respect to the more favorable profile identified in Cluster 2, entrepreneurs consistently interpreted high evaluations across dimensions as a form of validation. In some cases, this was associated with readiness for scaling, as positive feedback was seen as confirming both product value and market acceptance. In other cases, it was interpreted more cautiously as validation of the concept, still requiring further testing before broader commercialization. These findings suggest that even positive configurations do not eliminate uncertainty, but are instead incorporated into broader evaluative processes.

Ambiguity becomes particularly evident when entrepreneurs are confronted with conflicting signals within the same user feedback. Respondents frequently reported situations in which users recognized the value of a technology while simultaneously highlighting issues related to usability or comfort. In such cases, decision-making was described as iterative and contingent. One entrepreneur emphasized the need to assess whether usability constraints could limit market adoption, leading to internal discussions and case-by-case evaluations. Another highlighted the role of technical feasibility, noting that decisions often depend on whether design improvements are realistically achievable. A third respondent indicated that severe usability issues may require more radical product changes, including redesign, even when performance value is acknowledged.

These interpretations also reflect different approaches to managing trade-offs between competing product attributes. While some entrepreneurs emphasized the absence of fixed priorities, adapting decisions to perceived market needs, others described more strategic positioning choices, such as targeting specific niches when products exhibit strong performance but limited usability. In contrast, another respondent highlighted usability and comfort as primary decision criteria, even at the expense of other product attributes.

Concrete product decisions further illustrate how lead-user signals are translated into action. In one case, repeated user requests led to the addition of new features, reflecting a responsiveness to collaborative signals. In another, concerns about product size resulted in a deliberate trade-off, sacrificing some performance aspects to improve usability and compactness. In contrast, a different approach prioritized performance over design considerations, based on the assumption that usability and functionality outweigh aesthetic concerns.

These findings show that lead-user signal configurations do not provide direct prescriptions for action. Instead, they require interpretation through sensemaking processes that are shaped by entrepreneurs' assumptions, strategic priorities, and technical constraints. The same configuration of signals may thus lead to different decisions, reinforcing the view that lead-user inputs contribute not only to reducing uncertainty, but also to generating ambiguity in entrepreneurial decision-making.

5 Conclusions

This study examined how sport-tech entrepreneurs make sense of heterogeneous lead-user signals and translate them into product development decisions under conditions of uncertainty. The findings show that lead-user inputs do not necessarily provide clear and actionable guidance. Instead, they are structured in differentiated and sometimes conflicting configurations, which may generate ambiguity rather than reduce it.

By combining cluster analysis with qualitative insights, the study shows that the same configurations of user signals can be interpreted in different ways, leading to different product decisions. This extends lead-user theory by framing user inputs as signals that require interpretation, rather than as direct sources of innovation-relevant knowledge. It also contributes to research on entrepreneurial decision-making by highlighting the role of sensemaking in constructing coherence among divergent evaluations, as well as to paradox theory by showing how tensions between performance, usability, data reliability, and scalability are reflected in both user feedback and decision processes.

From a managerial perspective, the results suggest that entrepreneurs should not seek consensus in user feedback, but rather develop capabilities to interpret and prioritize heterogeneous signals. Ambivalent feedback, in particular, may reveal critical constraints and opportunities for product refinement.

The study is limited by the exploratory nature of the qualitative phase and the specific sport-tech context. Future research could extend the analysis to different settings and examine how interpretive processes evolve over time.

Overall, the study highlights that in technology-driven innovation, the key challenge is not only to access user feedback, but to make sense of it under conditions of ambiguity.

6 Ethical declaration

Ethical approval was not required for this study, as the research is based exclusively on secondary data and does not involve human participants, personal data, or any form of sensitive information.

7 AI declaration

No artificial intelligence (AI) tools were used in the preparation of this manuscript.

References

Bardus, M., Borgi, C., El-Harakeh, M., Gherbal, T., Kharroubi, S., & Fares, E. J. (2021). Exploring the use of mobile and wearable technology among university student athletes in Lebanon: a cross-sectional study. Sensors, 21(13), 4472.

Benar, N., Fatahi Milasi, S., & Nazarian, A. (2024). The behavioral intention of coaches in virtual reality: The roles of perceived usefulness, enjoyment, performance expectancy, self-efficacy, and ease of use. Journal of New Studies in Sport Management, 5(3), 1187-1198.

Brooke, J. (1996). SUS-A quick and dirty usability scale. Usability evaluation in industry, 189(194), 4-7.

Chamorro-Koc, M., Peake, J., Meek, A., & Manimont, G. (2021). Self-efficacy and trust in consumers' use of health-technologies devices for sports. Heliyon, 7(8).

Connelly, B. L., Certo, S. T., Ireland, R. D., & Reutzel, C. R. (2011). Signaling theory: A review and assessment. Journal of management, 37(1), 39-67.

Cusumano, M. A., Gawer, A., & Yoffie, D. B. (2021). Can self-regulation save digital platforms?. Industrial and Corporate Change, 30(5), 1259-1285.
Eppler, M. J., & Mengis, J. (2004). The concept of information overload: A review of literature from organization science, accounting, marketing, MIS, and related disciplines. The information society, 20(5), 325-344.
Henkens, B., Verleye, K., & Larivière, B. (2021). The smarter, the better?! Customer well-being, engagement, and perceptions in smart service systems. International Journal of Research in Marketing, 38(2), 425-447.
Iyer, E. S., Weinberg, A., & Bagot, R. C. (2022). Ambiguity and conflict: Dissecting uncertainty in decision-making. Behavioral Neuroscience, 136(1), 1.
Kim, T., Chiu, W., & Chow, M. K. F. (2019). Sport technology consumers: segmenting users of sports wearable devices based on technology readiness. Sport, Business and Management: An International Journal, 9(2), 134-145.
Kotlar, J., & Sieger, P. (2019). Bounded rationality and bounded reliability: A study of nonfamily managers' entrepreneurial behavior in family firms. Entrepreneurship theory and practice, 43(2), 251-273.
Lin, N., Wilden, R., Chirico, F., Ghasrodashti, E., & DeTienne, D. R. (2022). Persist or let it go: Do rational entrepreneurs make decisions rationally?. Journal of Business Venturing, 37(4), 106210.
Maitlis, S. (2005). The social processes of organizational sensemaking. Academy of management journal, 48(1), 21-49.
Poblete, L., Eriksson, E., Hellström, A., & Glennon, R. (2023). User involvement and value co-creation in well-being ecosystems. Journal of Health Organization and Management, 37(9), 34-55.
Porter, M. E., & Heppelmann, J. E. (2015). How smart, connected products are transforming companies. Harvard business review, 93(10), 96-114.
Prahalad, C. K., & Ramaswamy, V. (2004). Co-creation experiences: The next practice in value creation. Journal of interactive marketing, 18(3), 5-14.
Raff, S., Wentzel, D., & Obwegeser, N. (2020). Smart products: conceptual review, synthesis, and research directions. Journal of product innovation management, 37(5), 379-404.
Ramaswamy, V., & Ozcan, K. (2018). What is co-creation? An interactional creation framework and its implications for value creation. Journal of business research, 84, 196-205.
Ratten, V. (2024). Digital platform usage amongst female sport technology entrepreneurs. Journal of Small Business & Entrepreneurship, 36(5), 751-774.
Schmid, I., Wörner, J., & Leist, S. (2022). Automated identification of different lead users regarding the innovation process. Electronic Markets, 32(2), 945-970.
Simon, H. A. (2000). Bounded rationality in social science: Today and tomorrow. Mind & Society, 1(1), 25-39.
Smith, W. K., & Lewis, M. W. (2011). Toward a theory of paradox: A dynamic equilibrium model of organizing. Academy of management Review, 36(2), 381-403.
Tabesh, P., & Vera, D. M. (2020). Top managers' improvisational decision-making in crisis: a paradox perspective. Management Decision, 58(10), 2235-2256.
Troise, C., Ben-Hafaïedh, C., Tani, M., & Yablonsky, S. A. (2022). Guest editorial: new technologies and entrepreneurship: exploring entrepreneurial behavior in the digital transformation era. International Journal of Entrepreneurial Behavior & Research, 28(5), 1129-1137.
Venesz, B., Dőry, T., & Raišienė, A. G. (2022). Characteristics of lead users in different stages of the new product development process: A systematic review in the context of open innovation. Journal of open innovation: technology, market, and complexity, 8(1), 24.
Von Hippel, E. (1986). Lead users: a source of novel product concepts. Management science, 32(7), 791-805.
Weick, K. E. (1995). What theory is not, theorizing is. Administrative science quarterly, 40(3), 385-390.
Yin, C., Jiang, C., Jain, H. K., Liu, Y., & Chen, B. (2023). Capturing product/service improvement ideas from social media based on lead user theory. Journal of Product Innovation Management, 40(5), 630-656.
Yin, R. K. (2009). Case study research: Design and methods (Vol. 5). sage.
Zhang, T., Lu, C., Torres, E., & Cobanoglu, C. (2020). Value co-creation and technological progression: a critical review. European Business Review, 32(4), 687-707.

From Intellectual Capital Statements to AI-Enhanced Strategic Knowledge Management

Manfred Bornemann, Kay Alwert
Intangible Assets Consulting GmbH, Graz, Austria

Abstract

This paper explores the Intellectual Capital Statement (ICS) “made in Germany” as a participatory method for strategic management of intangible resources in knowledge-intensive organizations. The method emerged in the early 2000s and builds on the established categories of Human Capital, Structural Capital, and Relational Capital (Edvinsson and Malone, 1997; Sveiby, 1997; Stewart, 1997). It combines qualitative assessment with moderated workshops and cause-effect modelling (Alwert et al., 2013). ICS links the diagnosis of intangible resources with strategic fields of intervention and implementation measures (Bornemann et al., 2021).

ICS aligns with knowledge management standards such as ISO 30401 (ISO, 2023) and DIN SPEC 91443 (DIN, 2020). These frameworks define system requirements and governance structures. ICS contributes a process perspective that connects assessment with strategy development through moderated dialogue and shared interpretation, which supports organizational sensemaking (Weick, 1995).

A typical ICS application centers on a facilitated workshop. Participants assess mission-critical factors and justify their evaluations. This process externalizes tacit knowledge and supports its transformation into shared understanding (Nonaka and Takeuchi, 1995). The impact matrix visualizes interdependencies between elements of Intellectual Capital and organizational outcomes (Vester, 1984). Empirical evidence from more than 1,000 applications confirms usability and transferability (BMWi, 2010).

Recent applications show how AI supports the ICS process. In the case of SBR, a medium-sized accounting service firm undergoing generational transition and staff turnover, the 2025 ICS workshop was conducted in a hybrid format with automated transcription and GPT-based summarization. AI-supported documentation produced rapid and comprehensive output, reducing cognitive load and enabling deeper reflection on improvement measures In the case of ZM-I, anonymized workshop results were analyzed with large language models to generate draft recommendations, which were refined through structured dialogue. These applications indicate efficiency gains and enhanced analytical depth while interpretation remains anchored in organizational judgement.

ICS develops further in the context of AI-supported knowledge work. Concepts such as real-time dashboards and embedded ICS agents indicate a shift toward continuous observation of Intellectual Capital. These developments require governance structures that ensure transparency and contextual sensitivity. ICS retains its role as a participatory method for organizational sensemaking and strategic transformation (Weick, 1995).

Keywords: Intellectual Capital Statement, Knowledge Management; Intangible Assets; Artificial Intelligence; Participatory Methods

Paper type: Academic Research Paper

1 Introduction

More than 30 years into the knowledge society (Drucker, 1993), at the dawn of the “age of artificial intelligence” (Bostrom, 2014), where global economies are shaped more by expertise, networks, and intangible assets than by physical goods, organizations still struggle to make the invisible visible. Strategic decisions are often based on financial metrics, historical structures, standard recipes, and individual preferences of top managers, while the dynamics caused by Intellectual Capital (IC) remain underexplored and undervalued.

1.1 Intellectual capital statements: origins and ongoing relevance

The Intellectual Capital Statement (ICS) "made in Germany" emerged as a response to this gap. Developed in the early 2000s (Alwert et al 2004 und 2013, Mertins et al 2005), it offered a practice-oriented framework to assess and develop intangible resources through an interplay of measurement, structured dialogue, participatory workshops, and visual modelling of cause-and-effect networks and has always emphasized strategic relevance and internal resonance. It helped anchor knowledge management within broader organizational development and positioned Intellectual Capital as a driver of competitive advantage.

Two decades later, ICS is still dealing with a central dilemma: How can organisations capture their collective knowledge in a credible, comprehensive, and understandable way so that it can be used for strategic management as well as for external accountability?

This question has been addressed at a normative level by recent standards for knowledge management, including ISO 30401:2018, revised in 2023, and the DIN SPEC 91443:2020 (see: DIN, 2020; ISO, 2023). Both standards provide systematic guidance for managing intangible resources, particularly for knowledge-intensive and small-to-medium-sized enterprises. While ISO 30401 establishes global principles for structured KM systems, DIN SPEC 91443 focuses on context-sensitive implementation for SMEs. The ICS, however, stands out as the only method explicitly integrating assessment and development: it identifies periodic changes and prioritizes fields of intervention, and systematically supports the implementation processes necessary to enact these changes (Bornemann et al, 2021).

1.2 Technological changes offer new opportunities

We are entering a new phase in the evolution of management: the dawning age of artificial intelligence (Brynjolfsson & McAfee, 2017; Tegmark, 2018). Language models, real-time data analytics, and predictive tools offer impressive possibilities to complement human judgment, track knowledge flows, and simulate organizational dynamics.

This article takes ICS as both a method and a perspective. We review its conceptual architecture and operational logic (Chapter 2), examine current practice through two updated case studies (Chapter 3), and explore how AI may reshape its future applications (Chapter 4). We explore how the method can remain relevant in a digitally transformed world, and whether ICS can evolve from a periodic assessment tool to a dynamic, learning-oriented system that supports strategic reflection in real time.

Our argument builds on earlier findings (Bornemann et.al, 2021) but moves beyond them. We show that ICS continues to offer value in conventional organizational settings and provides a powerful frame to reflect on the quality of knowledge governance in AI-supported environments. Case experiences suggest that AI can reduce the cognitive burden of documentation and enhance the precision of ICS outputs, without replacing the participatory essence of the method. In fact, by automating routine tasks and surfacing latent insights, AI may enable richer reflection, faster learning, and deeper insights.

More broadly, we consider ICS a potential bridge between data-driven and sensemaking practices; between planned strategy and lived behaviour; and between documentation and transformation. In an era that risks reducing knowledge to quantifiable artifacts, the ICS method reminds us that knowledge is also a social process, embedded in language and context, shaped through dialogue, and enacted in practice (Latour 2005). As such, it must be made explicit not only for reporting, but for understanding and improvement.

2 The Intellectual Capital Statement (ICS) Methodology

The Intellectual Capital Statement "made in Germany" is a structured approach to assess, visualize, and manage intangible resources. It is grounded in the theoretical frameworks of IC management and knowledge-based value creation. Drawing on the works of Edvinsson and Malone (1997), Sveiby (1997), and Stewart (1997), ICS expands the triadic structure of IC - Human, Structural, and Relational Capital - into a pragmatic, operational method. Within this triad organizational strengths, weaknesses, and development needs are analysed. Specific drivers of IC include e.g. employee competence, leadership, innovation & knowledge processes, organisational culture, IT infrastructure, external partnerships, and customer relations (Bornemann and Reinhardt, 2017).

The ICS methodology is built around a structured, participatory, and workshop-based process. At its core is a consensus-driven assessment of qualitative indicators to identify strengths and weaknesses, as well as a visual cause-effect model that links IC resources to value creation processes and strategic objectives. This approach ensures alignment with organizational priorities and stakeholder expectations.

To support this process, a dedicated software tool - the ICS Toolbox (BMWi, 2013) - has been developed. It standardizes documentation, supports strategic management, and ensures methodological consistency. The toolbox guides the process from preparation and team selection, through data collection and assessment, to the final report.

A typical ICS implementation includes the following eight steps: (1) preparatory phase with team composition, goal setting an reflection of the business model, (2) data collection from available sources, (3-5) the ICS workshop to assess the Intellectual Capital and its interplay with processes and goals, which is the core event moderated by a trained facilitator, (6) diagnosis with validation of results and feedback loops, (7) deduction of concrete actions for improvement, (8) documentation, reporting & communication. During the workshop, participants from various hierarchical levels and functions assess a set of mission-critical factors using a scale from 0% to 120%. Crucially, each assessment must be justified through structured dialogue and documented accordingly, often with narrative-rich explanations.

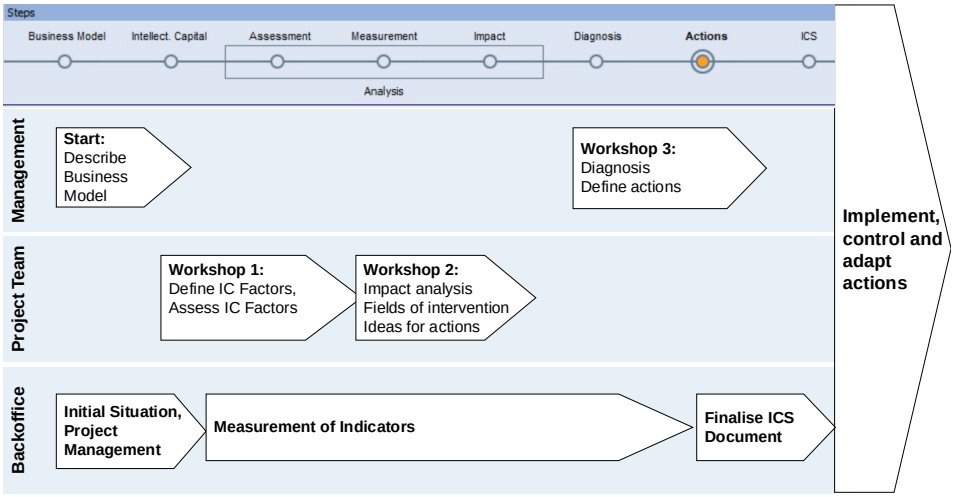


Figure 1 The implementation process of am ICS (Bornemann et al 2021)

This dialogic approach serves the following purposes: it fosters shared understanding across departments, surfaces tacit knowledge, and builds a common language for discussing intangibles. ICS is both, a reporting mechanism, and an organizational development tool. It acts as a mirror, enabling reflection and self-assessment through systemic feedback (Bornemann and Reinhardt, 2017).

A central feature of ICS is its integration of strategic goals with intangible resource assessment. This contrasts with traditional balance sheet analogues or capital inventory lists. ICS offers dynamic insights into how knowledge resources contribute to performance, agility, and long-term value creation. This is particularly relevant to all organisations, where strategic capacity and transparency are essential (Will, 2007).

The method emphasizes inclusiveness and participation, often drawing from organizational development theory (Weick, 1995). Employees become co-creators of the ICS narrative, thereby enhancing legitimacy and implementation readiness. This participatory ethos mitigates resistance and enables bottom-up commitment to actions for change. By promoting systemic thinking, ICS contributes to organizational learning and resilience (Senge, 1990).

ICS has been applied in over 1,000 cases in Germany, including manufacturing firms, public administration bodies, hospitals, and research institutions. Evaluation studies funded by the German Federal Ministry of Economics (BMWi) have confirmed the method's usability, impact, and transferability BMWi (2010).

In recent years, ICS has shown potential for adaptation in the context of digitalization. Emerging tools for automated transcription, natural language processing, and semantic analysis can support the documentation process. Likewise, AI-supported summarization and classification routines offer new ways to reduce workload and enhance consistency. These innovations are particularly relevant for ICS implementations in remote or hybrid environments.

As outlined in Bornemann, Alwert, and Will (2021), the ICS offers a unique blend of strategic clarity and operational relevance. Its structured yet adaptable design makes it particularly suitable for knowledge-intensive organizations seeking to align intangible assets with long-term value creation.

3 Reflections on Case Applications of ICS using AI-Support

The evolution of the Intellectual Capital Statement method reflects its theoretical robustness as well as its adaptability to emerging technological contexts. While original ICS practice emphasized, strategic alignment, structured dialogue and consensus-based assessments, recent implementations show how a digital focus within an ICS stimulates strategy development and how digital tools - particularly AI-assisted transcription and summarization - can meaningfully enhance the ICS process. This chapter discusses two organizations previously studied in the 2021 Journal of Intellectual Capital article (Bornemann et al, 2021) and explores their latest experiences.

3.1 Case Study: SBR (2025)

SBR, a medium-sized accounting service firm with more than 20 years of ICS experience, returned to the methodology in 2025 to reflect on recent organizational changes. While its historical ICS efforts were driven by

quality management and operational transparency (Bornemann et al., 2021), the current context introduced new dynamics: a transition in leadership to the next generation and a phase of higher-than-usual staff turnover. The ICS served as both: a retrospective reflection tool and a structured onboarding mechanism for the incoming CEO, allowing insights into the company's historical evolution and team perspectives.

For the first time, the workshop was conducted in a hybrid format via Microsoft Teams, with one remote participant. Full-session recording and automated transcription were employed - with prior consent from participants - to support documentation. Two parallel documentation strategies were evaluated:

Manual ICS Documentation: A student in her final bachelor's year documented the assessments in the ICS Toolbox, following established procedures with coaching support.

AI-Supported Documentation: The session transcript was processed using a tailored GPT model trained on ICS guidelines and reference documents. It generated summaries of IC drivers, synthesized rationales, and proposed improvement measures.

Both documentation types were available for the second workshop, where improvement strategies were discussed. Comparatively, the AI-generated report was faster, richer in detail, and more accessible in tone, while the manually prepared document, though delayed, offered precision and internal familiarity (e.g. local wording or context markers). Participants noted a reduction in stress and cognitive load through AI support, particularly during intensive sessions.

The AI model also produced a structured proposal of improvement measures, which were reviewed, discussed, and adapted by the team. This demonstrates that AI can complement human judgment. And it supported the organization's cultural commitment to transparency and engagement. The CEO endorsed the hybrid method, recognizing its potential for process efficiency and deeper insights.

3.2 Case Study: ZM-I (2025)

ZM-I group is a leading German engineering company with approximately 140 knowledge workers in seven companies. The firm plans, tests and assesses the structural frameworks of a wide range of building types as a highly knowledge intensive technical B2B service. As a medium-sized enterprise with increasingly decentralised structures, ZM-I has relied for over a decade on the Intellectual Capital Statement – Made in Germany methodology to systematically identify, steer and develop its intangible resources and manage organisational complexity.

In 2025, the IC management process was used to support a strategic reorientation in the context of the company's ongoing digital transformation. A central focus was the expansion of the digitalisation strategy initiated in 2022 through the integration of a company-wide AI strategy, aimed at aligning future developments with the demands of data-driven engineering services.

The analysis phase addressed essential technological structures, cross-location knowledge exchange and the competencies required to realise the extended strategy. The objective was to assess the organisational prerequisites for successfully implementing both digital and AI-driven initiatives.

Another advancement in the recent ICS process involved leveraging large language models (LLM) to analyze anonymized workshop outcomes, formulate summaries and generate recommendations for actions. These AI-driven suggestions were subsequently reviewed by management as well as workshop participants and incorporated as valuable input for refining the final action plan.

4 Future Scenarios and Systemic Learning in ICS Applications

Building on the adaptive and dialogic qualities of the ICS, recent discussions have begun to envision future scenarios where real-time feedback systems, augmented by artificial intelligence, transform the ICS from a periodic reflection tool into a continuous strategic monitoring interface. These developments open possibilities for integrating ICS logic into operational dashboards, supporting management with dynamically updated insights into the status of intangible resources and their relation to business outcomes.

4.1 Toward a real-time ICS dashboard

One scenario suggests a dedicated "ICS Agent" embedded within enterprise information systems. Drawing on data streams from collaborative platforms (e.g., Microsoft Teams, Google Workplace, or others), HR systems, knowledge repositories, and project management tools, the ICS Agent could autonomously collect, interpret, and visualize relevant data and indicators.

This agent would complement workshops and maintain organizational awareness between formal ICS cycles. The agent could detect emerging patterns of stress or capacity, highlight over- or underutilized competencies,

or raise alerts when strategic knowledge becomes siloed or is in risk to get lost. In this role, the ICS agent supports resilience and anticipatory learning, contributing to what Nonaka and Takeuchi (1995) describe as a knowledge-creating company.

Echoing broader trends in digital twin systems and predictive analytics, this approach envisages ICS as an integral component of the real-time management cockpit. It retains the reflective and participatory nature at the heart of ICS m.i.G., ensuring that dynamic oversight does not come at the expense of meaningful engagement.

An additional innovation concerns internal communication: Instead of distributing a final report in static form, an agentic approach could allow employees to explore results, action plans, and IC relations through dialogue. This format follows the principles of user-centred, conversational knowledge delivery, aligned with recent developments in digital workplace environments.

4.2 The ICS impact matrix as learning catalyst

Another line of development focuses on the ICS impact matrix, a cause-effect model, inspired by Frederic Vester's sensitivity model (Vester, 1984), that links intangible resources to business processes and outcomes. Historically, this matrix has been constructed through moderated team workshops, based on perceived and expected influences of IC elements.

Future systems may allow a comparison between these designed models and empirical usage patterns derived from actual process data, communication logs, and performance metrics. Such comparisons would reveal where documented logic diverges from lived organizational behaviour. This new dual-layered approach could offer two learning paths:

Adaptive Realignment: If the empirical data reflect a coherent operational logic, the designed impact model might be revised to better match actual practices, enhancing strategic realism.

Reflective Reconciliation: If the data reveal performance gaps or dysfunctional patterns, this discrepancy becomes a stimulus for dialogue and change. ICS teams can review the designed model, question misalignments, and co-create strategies to align behaviour with intention.

In both cases, the ICS evolves from a descriptive document into a generative system that surfaces contradictions and supports learning processes.

4.3 From bureaucratic routine to organizational learning

ICS has always positioned itself between strategic alignment and participatory sensemaking. Embracing AI for efficiency and for deeper diagnosis, allows organizations to turn ICS into a catalyst for transformative learning, a process that challenges individual and organizational assumptions, stimulates knowledge exchange, fosters creative responses to complex challenges, and promotes proactive change (Nonaka & Takeuchi, 1995; Senge, 1990).

This potential depends on governance choices, digital literacy, and the degree to which ICS remains anchored in shared values and collective interpretation. Technological augmentation should reinforce ICS's core purpose: to make intangible resources visible, discussable, and improvable (ISO, 2023).

Future ICS applications may include AI-supported plausibility checks across assessments. LLMs can compare argument structures underlying evaluations of Human Capital, Structural Capital, and Relational Capital and identify inconsistencies in reasoning. Similar patterns of deficits with comparable strategic relevance can be highlighted when they lead to diverging scores. Such analysis does not alter the assessment itself but supports participants in refining their justifications and aligning evaluation logic. This strengthens internal coherence and increases transparency of the ICS process.

5 Discussion and Outlook

The integration of AI into the ICS process is raises fundamental issues about the nature of knowledge work, organizational learning, and the trustworthiness of decision-making tools. As our case studies and conceptual reflections show, AI has the potential to both accelerate and deepen ICS processes, but not without new risks and unresolved challenges. In this concluding chapter, we synthesize key findings, assess opportunities and limitations, and outline implications for practice.

5.1 Efficiency gains and quality improvements

AI-supported ICS processes clearly demonstrate potential for improving efficiency. Automated transcription, summarization, and pattern recognition, especially when applied to recorded workshop dialogues, can significantly reduce the time and cognitive load required for ICS documentation. This is supported in the SBR case, where a tailored GPT model produced actionable summaries that were available the day after the ICS workshop. Beyond speed, quality improvements were also noted: participants found the AI-supported output more readable and comprehensive than traditional notes. In addition, the facilitator could focus on group dynamics rather than simultaneous documentation.

In the case of ZM-I, the interplay between ICS and AI contributed meaningful to strategy formulation, communication, and the development of enhanced improvement measures.

In future applications, AI may also enable deeper learning within ICS by linking structured outputs to behavioural data. One of the core elements of ICS is the impact matrix, which visualizes how specific IC components (e.g., staff development or IT infrastructure) affect defined business outcomes. These matrices could be validated - or challenged - by behavioural data drawn from process mining, digital workflows, or communication data. The comparison between designed models and observed practice opens another space for systemic learning and qualitative improvement. The obvious prerequisite is the explicit agreement of employees to this procedure.

Combining the ICS method with a well-trained AI agent, AI can help structure unfiltered dialogue into actionable ICS components such as drivers, impact relationships, and measures. It also can help to check if the intended and lived organisation diverge significantly. While these opportunities do not replace the need for critical review and ethical judgment, they change the balance between human interpretation and technical assistance.

5.2 Risks: distortion, bias, and convergence

Integrating AI into ICS workflows introduces risks that affect how assessments are constructed and interpreted. LLMs structure outputs along with statistical regularities, which can distort context-rich discussions and shift the emphasis of arguments. These effects influence the representation of relationships between elements of Intellectual Capital and can alter the perceived relevance of drivers.

A structural risk lies in convergence toward dominant patterns. This dynamic reduces variance in assessments and weakens the differentiation of strategic positions. ICS depends on identifying organization-specific constellations of Human Capital, Structural Capital, and Relational Capital. Convergence within AI-supported outputs can obscure these distinctions.

Bias operates across training data, prompt design, and organizational context. These influences shape how arguments are framed and how drivers are weighted within the assessment. In evaluations that relate to internal structures and performance, such effects can influence acceptance and interpretation of results.

AI-supported consistency checks introduce a secondary dynamic. Comparisons across assessments can align evaluations that differ in context or strategic relevance. Divergence in assessments often reflects relevant organizational tensions and contributes to diagnostic depth. ICS benefits from preserving these differences as part of the reflection process.

5.3 Ethical considerations and the role of AI

While AI can support the ICS process, it should not assume authority over normative judgments, such as interpreting commitment, assessing employees, or recommending cultural shifts. ICS remains a participatory tool; the danger lies in reducing it to a mechanical procedure that outputs texts with pseudo-objectivity.

A useful analogy can be drawn from education: Some researchers have compared machine learning with learning and teaching processes of humans (Moravec 1988, Tegmark 2018, Merlo & Alwert 2025). In this perspective, AI behaves like a "learning actor" rather than a mere instrument. This suggests that organizations should treat AI with the same attentiveness and respect they would give to onboarding a junior team member (Merlo & Alwert 2025): clarify expectations, training, monitor learning progress, and define boundaries of responsibility.

5.4 Outlook: a cyclical, AI-enhanced ICS

Looking ahead, the ICS could evolve from an episodic workshop model into a continuous, AI-supported knowledge management system supporting ISO 30401. Real-time ICS dashboards, intelligent agents trained on company-specific ICS profiles, and integration with strategic planning cycles are conceivable. In such scenarios,

ICS serves as documentation tool as well as a dynamic learning environment for the whole organisation and its strategic development.

Future ICS iterations may rely less on verbal workshops and more on synthesized observations from digital behaviour, yet their interpretive power will still depend on human dialogue. The ICS may thus serve as a meeting point between empirical data, shared sensemaking, and strategic transformation.

6 Acknowledgment on the Use of Artificial Intelligence

During the preparation of this article, the authors received support from artificial intelligence in selected stages of the writing process. ChatGPT-5.4 (OpenAI) were used to assist in the formulation and refinement of text passages based on preliminary drafts. ChatGPT-5.4 as well as deepL translate & write were employed to support the translation of sections originally written in German, and to improve the clarity, consistency, and readability of the manuscript, as English is not the authors' first language. ChatGPT-5.4 was also employed to assist in the formatting of references in APA style. All AI-assisted contributions were carefully reviewed, edited, and validated by the authors. The authors accept full responsibility for the accuracy, interpretation, and presentation of all content included in this article.

7 Ethics Declaration

This study did not involve experiments on human participants or animals requiring ethical approval. The case studies presented are based on organizational data and publicly available information. Where organizational or individual data were used, they were included with the explicit consent of the respective organizations. No personal or sensitive data were processed in a manner that would require formal ethical review. The authors declare no conflict of interest. No external funding was received for this study.

References

Alwert, K., Bornemann, M., & Kivikas, M. (2004). *Leitfaden Wissensbilanz – Made in Germany*. Bundesministerium für Wirtschaft und Arbeit (BMWA).

Alwert, K., Bornemann, M., Will, M., & Wuscher, S. (2013). Wissensbilanz – Made in Germany: Leitfaden 2.0. Bundesministerium für Wirtschaft und Technologie (BMWi).

Bornemann, M., Alwert, K., & Will, M. (2021). Lessons learned in intellectual capital management in Germany between 2000 and 2020 – History, applications, outlook. *Journal of Intellectual Capital*, *22*(3), 560–586. https://doi.org/10.1108/JIC-03-2020-0085

Bornemann, M., & Reinhardt, R. (2017). Handbuch Wissensbilanz: Umsetzung und Fallstudien (2nd ed.). Erich Schmidt Verlag. https://doi.org/10.37307/b.978-3-503-17441-6

Bundesministerium für Wirtschaft und Technologie (BMWi). (2010). *Vorstudie "Wissensstandort Deutschland"*. BMWi.

Bundesministerium für Wirtschaft und Technologie (BMWi). (2013). *Wissensbilanz Toolbox 2.0: Software zur Erstellung einer Wissensbilanz*.

Bostrom, N. (2014). *Superintelligence: Paths, dangers, strategies*. Oxford University Press.

Brynjolfsson, E., & McAfee, A. (2017). *Machine, platform, crowd: Harnessing our digital future*. W. W. Norton & Company.

Edvinsson, L., & Malone, M. S. (1997). *Intellectual capital: Realizing your company's true value by finding its hidden brainpower*. Harper Business.

Latour, B. (2005). *Reassembling the social: An introduction to actor-network theory*. Oxford University Press.

Merlo, T. R., & Alwert, K. (2025). Ethical governance of generative AI: Accountability, oversight, and societal impact. In *AI-driven revolution transforming the business landscape*. World Scientific. https://doi.org/10.1142/q0488

Mertins, K., Alwert, K., & Heisig, P. (2005). *Wissensbilanzen: Intellektuelles Kapital erfolgreich nutzen und entwickeln*. Springer.

Moravec, H. (1988). *Mind children: The future of robot and human intelligence*. Harvard University Press.

Nonaka, I., & Takeuchi, H. (1995). *The knowledge-creating company: How Japanese companies create the dynamics of innovation*. Oxford University Press.

Senge, P. M. (1990). *The fifth discipline: The art and practice of the learning organization*. Currency.

Stewart, T. A. (1997). *Intellectual capital: The new wealth of organizations*. Doubleday.

Sveiby, K. E. (1997). *The new organizational wealth: Managing and measuring knowledge-based assets*. Berrett-Koehler.

Tegmark, M. (2018). *Life 3.0: Being human in the age of artificial intelligence*. Penguin Books.

Vester, F. (1984). *Die Kunst, vernetzt zu denken: Ideen und Werkzeuge für einen neuen Umgang mit Komplexität*. Deutsche Verlags-Anstalt.

Weick, K. E. (1995). *Sensemaking in Organizations*. Sage Publications.

Will, M. (2007). Wissensbilanzen als Instrumente der Selbststeuerung und -beobachtung. In G. Reinmann & H. Mandl (Eds.), *Wissensmanagement in der beruflichen Bildung* (pp. 243–256). Waxmann.

Innovation Orchestration through Innovation Labs: Evidence from Intellimech

Francesco Santarsiero[1], Giovanni Schiuma[1], Alessandro Massaro[1], Achille Claudio Garavelli[2]

[1] Università LUM Giuseppe Degennaro, Casamassima, Italy

[2] Politecnico di Bari, Bari, Italy

Abstract

Innovation labs are increasingly used to connect firms, universities, technology providers and local institutions around experimentation, knowledge sharing and digital transformation. However, less is known about how innovation labs orchestrate innovation when they operate not as isolated corporate laboratories, but as membership-based collaborative infrastructures serving a network of firms. This paper investigates this issue through an exploratory qualitative case study of Intellimech, an Italian private consortium for mechatronics and industrial innovation located in the Kilometro Rosso innovation district in Bergamo. Drawing on publicly available documentary evidence, including Intellimech's institutional webpages, statute, project descriptions and published case histories, the paper conceptualizes Intellimech as an innovation lab and intermediary that transforms distributed technological needs into shared learning opportunities, pre-competitive projects, open innovation initiatives and firm-level experimentation. The findings identify five orchestration mechanisms: need sensing and challenge framing; shared knowledge creation and absorptive capacity building; pre-competitive experimentation; boundary spanning across firms, universities and institutional actors; and translation of shared knowledge into firm-specific pilots and organizational routines. Examples such as AIxLAB, the Knowledge Management Cluster, Watchman, Harmony, Shop4cf/Warning and other mechatronics projects show how the lab logic can reduce the cost and uncertainty of innovation for firms while supporting knowledge circulation within a territorial ecosystem. The paper contributes to research on innovation labs, open innovation and knowledge management by showing that innovation orchestration is not limited to coordinating technological projects, but also involves the design of social, cognitive and organizational conditions that allow firms to learn collectively and innovate individually. The study also provides managerial implications for local ecosystems that seek to build sustainable innovation capacity through shared research, mentoring, training, technology scouting and pilot projects. Limitations concern the reliance on secondary public sources; future research should integrate interviews with Intellimech managers and member companies.

Keywords: innovation labs; innovation orchestration; open innovation; knowledge management; digital transformation

Paper type: Academic research paper

1 Introduction

Firms increasingly innovate in conditions of technological turbulence, knowledge dispersion and rising uncertainty. In advanced manufacturing and mechatronics, digital transformation is not merely a matter of adopting individual technologies such as artificial intelligence, robotics, industrial data platforms or sensor-based systems. It requires the capacity to identify relevant technological trajectories, translate them into business problems, mobilize complementary expertise and test solutions in realistic contexts. For many small and medium-sized enterprises, and even for larger firms facing fast technological change, this process is difficult to manage internally because knowledge is distributed across multiple actors and because experimentation involves financial, technical and organizational risks.

Against this background, innovation labs have emerged as organizational arrangements designed to support experimentation, learning and the development of new solutions. They are often discussed as physical or virtual spaces where firms, users, universities and technology providers jointly explore opportunities, prototype solutions and accelerate innovation. Yet the literature still tends to focus on innovation labs as spaces, methods or project containers, while less attention is paid to their role as orchestrators of innovation ecosystems (Schiuma and Santarsiero, 2024). This is particularly relevant when the lab does not belong to a single firm, but works as a shared infrastructure for a network of companies that contribute resources and expect access to knowledge, training, mentoring, technology scouting, open innovation opportunities and pilot projects.

This paper addresses the following research question: how can an innovation lab orchestrate innovation and knowledge processes within a network of firms and local ecosystem actors? To answer this question, the paper analyses Intellimech, a private Italian consortium for mechatronics based in Bergamo. Intellimech was founded

in 2007 and presents itself as a private consortium of large, medium-sized and small firms focused on interdisciplinary research in mechatronics for different industrial sectors. It operates at the intersection of companies, universities, technology providers and institutional actors, offering shared research, technology scouting, research and development, consulting, training and project-based experimentation (Intellimech, 2026).

The relevance of Intellimech lies in its hybrid nature. It is not simply a consultancy provider, nor merely an R&D contractor. It is a membership-based innovation infrastructure in which firms enter a network and contribute financially to a common organization that can identify needs, aggregate demand, reduce the individual cost of technological exploration and transform scattered expertise into shared innovation pathways. The statute of the consortium refers to an initial participation fee and to an annual management quota paid by members to cover operating and management costs (Intellimech, 2019). This governance logic is consistent with the idea of an innovation alliance: companies invest in a shared organizational platform that gives access to knowledge, projects and connections that would be harder to develop alone.

The paper contributes to the IFKAD debate by linking innovation labs to knowledge management and innovation orchestration. It argues that the value of a lab such as Intellimech is not only in the technologies it develops, but also in its capacity to organize knowledge flows, activate trust among heterogeneous actors, support absorptive capacity and convert exploratory activities into actionable innovation projects. The paper is structured as follows. The next section discusses the theoretical background on innovation labs, open innovation, knowledge management and orchestration. The methodology then explains the exploratory case study based on public documentary evidence. The empirical section presents Intellimech and identifies the main orchestration mechanisms. The discussion develops a conceptual model of innovation orchestration through labs and highlights theoretical and practical implications.

2 Theoretical Background

Innovation increasingly depends on the ability to connect heterogeneous knowledge sources. Open innovation theory suggests that firms should combine internal and external knowledge to accelerate innovation processes and access ideas, technologies and capabilities beyond organizational boundaries (Chesbrough, 2003; West and Bogers, 2014). This assumption is particularly relevant in complex technological domains, where no single organization possesses all the knowledge required to develop and implement new solutions. However, openness does not automatically generate innovation. External knowledge must be identified, interpreted, selected, recombined and translated into organizational action. This makes knowledge management a central component of open innovation.

From a knowledge management perspective, innovation depends on the creation, sharing, codification and application of knowledge (Nonaka and Takeuchi, 1995; Spender, 1996). In inter-organizational settings, knowledge processes are even more complex because they involve differences in language, routines, goals and absorptive capacity. Firms need mechanisms that make knowledge visible and actionable, such as workshops, assessment tools, pilot projects, roadmaps, communities of practice and shared repositories. These mechanisms can be interpreted as boundary objects: artefacts that allow actors with different backgrounds to collaborate without requiring full consensus on every detail.

Innovation labs are one way to institutionalize these mechanisms. Previous research has described innovation labs as spaces and organizational arrangements that support creativity, experimentation, prototyping and learning (Lewis and Moultrie, 2005; Magadley and Birdi, 2009). They typically combine physical infrastructure, facilitation methods, technical expertise and project-based experimentation. In their most mature form, innovation labs do not only generate ideas; they enable organizations to move from ideation to proof-of-concept, from proof-of-concept to adoption, and from adoption to organizational learning (Santarsiero et al., 2022).

The orchestration perspective helps explain how this happens. Innovation orchestration refers to the set of activities through which a focal actor mobilizes and coordinates a network of participants around innovation opportunities (Dhanaraj and Parkhe, 2006; Nambisan and Sawhney, 2011). Orchestration involves identifying opportunities, attracting partners, shaping roles, managing knowledge flows, building trust and supporting the integration of complementary resources. In innovation ecosystems, the orchestrator does not necessarily own all resources; rather, it creates the conditions under which distributed resources can be combined.

This paper integrates these streams by conceptualizing an innovation lab as an orchestration infrastructure. In this view, the lab is not only a place where innovation projects occur, but a mechanism for sensing needs, translating technological trends, building shared understanding, coordinating experiments, and supporting the diffusion of knowledge across firms. Such a view is particularly useful for membership-based labs and consortia,

where companies join a common platform to access capabilities, reduce uncertainty and participate in pre-competitive innovation activities.

3 Methodology

The paper adopts an exploratory qualitative case study design. The case study method is appropriate because the research question concerns how innovation orchestration unfolds in a real organizational and ecosystem context. The selected case, Intellimech, is theoretically relevant because it combines characteristics of an innovation lab, a private consortium, an open innovation intermediary and a territorial knowledge platform.

The empirical material consists of public documentary evidence collected from Intellimech's official website and related public documents. The evidence includes institutional pages describing the consortium, its method, technological areas, member network, research sharing, technology scouting, R&D activities, consultancy, connections with universities and institutions, and its location within the Kilometro Rosso innovation district. It also includes project and case-history pages concerning AIxLAB, the Knowledge Management Cluster, Watchman, Harmony, Shop4cf/Warning, robot-oriented applications and other mechatronics projects. The statute of the consortium was used to understand the formal governance logic of membership, participation fees and annual management quotas.

The analysis followed a qualitative document analysis procedure. First, the material was read to identify descriptive elements concerning Intellimech's mission, services, governance and ecosystem position. Second, statements and examples were coded according to innovation orchestration activities: sensing needs, selecting challenges, sharing knowledge, connecting actors, experimenting, piloting, translating solutions and supporting adoption. Third, the codes were grouped into broader orchestration mechanisms. The resulting interpretation is not intended to provide a complete internal evaluation of Intellimech, because no interviews or proprietary data were available at this stage. Rather, the paper offers an analytically grounded reading of Intellimech as a public-facing case of innovation orchestration through a lab-consortium model.

This methodological limitation also defines the contribution of the paper. Instead of claiming causal evidence, the study develops a conceptual and illustrative account that can be strengthened in future research through interviews with Intellimech managers, technical staff and member firms.

4 Case Background: Intellimech as a Lab-Consortium for Industrial Innovation

Intellimech defines itself as a private consortium of companies focused on interdisciplinary research in mechatronics (Intellimech, 2026). Its public presentation emphasizes the participation of large, medium-sized and small firms and its role in supporting applications across different industrial sectors. The consortium is located in the Kilometro Rosso innovation district in Bergamo, described as a meeting point between research and business that fosters synergies among companies, research centres, laboratories, professional services and higher education. Intellimech also has a Smart Lab at the POINT technological innovation hub in Dalmine.

The member network is heterogeneous. The public list of members includes companies from automation, electronics, energy, ICT, mechanics, construction and services, with firms such as ABB, Brembo, Dassault Systemes, Dmg Mori, Evoca Group, Fassi, Gefran, Kilometro Rosso, Lovato Electric, Oròbix, Phoenix Informatica, Siemens, SIAD Macchine Impianti, Sorint, Tesmec and many others (Intellimech, 2026). This heterogeneity is important because it creates the conditions for cross-sectoral learning. Firms with different technological backgrounds and market positions can encounter similar challenges in data management, automation, AI adoption, system integration, quality control, remote diagnostics, predictive maintenance and organizational knowledge management.

The Intellimech method starts from the needs of members and aims to transform them into innovative technological solutions. Its public description highlights problem solving, access to latest research, technology scouting and the definition of pre-competitive projects tailored to members' needs. The consortium also emphasizes the creation of a knowledge network based on cooperation, constructive confrontation and the exchange of ideas and strategies. This is central to the interpretation proposed in this paper: Intellimech works as an innovation lab because it provides structured processes through which firms can move from needs to projects, from technological trends to experimentation, and from individual problems to shared learning.

The value proposition can be read as a portfolio of orchestration services. First, Intellimech promotes shared research projects on key mechatronics topics and supports the development of pre-competitive platforms for member companies (Intellimech, 2026). Second, it carries out technology scouting to identify themes that are strategically relevant for firms and translate them into specific industrial needs. Third, it develops prototypes and manages research projects commissioned by members or third-party companies. Fourth, it participates in regional, national and European funding schemes, involving members as pilot cases. Fifth, it offers customized

consulting in mechatronics and industrial technologies. Finally, it supports training and continuous updating of technicians, which is particularly relevant in fast-changing technological domains.

This combination of activities suggests that Intellimech is more than a project provider. It can be understood as a membership-based innovation lab that transforms the economic contribution of members into shared access to knowledge, experimentation and ecosystem connections. The annual management quota described in the statute supports the functioning of the consortium, while additional projects and funded initiatives create opportunities for more focused experimentation (Intellimech, 2019). In this sense, Intellimech operates as an innovation alliance: firms do not simply buy isolated services, but participate in a collective infrastructure that can reduce uncertainty and increase innovation capacity.

5 Findings: Orchestration Mechanisms

The analysis identifies five main mechanisms through which Intellimech appears to orchestrate innovation: need sensing and challenge framing; shared knowledge creation; pre-competitive experimentation; boundary spanning; and translation into firm-level pilots and routines.

Table 1. Public evidence and orchestration mechanisms in the Intellimech case

Public evidence	Examples	Orchestration mechanism	Knowledge effect
Method and technology scouting	Needs of members; scouting of relevant themes; definition of pre-competitive projects	Need sensing and challenge framing	Transforms technological trends into shared industrial problems
Research sharing	Shared R&D projects and pre-competitive platforms	Collective agenda setting	Reduces duplication and creates common learning trajectories
R&D and funded projects	WATCHMAN, VIPS, SHOP4CF/WARNING, HARMONY	Pre-competitive experimentation	Creates demonstrators, prototypes and evidence for adoption
AIxLAB	Lovato Electric and COES experiences	Capability building and application support	Moves AI from abstract interest to firm-specific use cases
Knowledge Management Cluster	Phoenix, SIAD MI, Tesmec and other firms	Shared knowledge creation	Supports codification, organizational memory and peer learning
Kilometro Rosso and university links	Universities, research centres, institutional actors and innovation district	Boundary spanning	Combines scientific knowledge, industrial needs and territorial legitimacy

5.1 Need sensing and challenge framing

The first orchestration mechanism concerns the identification and framing of members' needs. Intellimech states that it starts from the needs of members and that the early phases of its operational approach involve technology scouting on themes considered relevant by companies. Meetings with firms are described as fundamental for selecting shared projects, because they allow broad technological domains to be translated into specific industrial requirements.

This mechanism is important because technological trends often remain too abstract for firms. AI, robotics or industrial data spaces become valuable only when they are framed in relation to concrete problems such as defect detection, quality monitoring, maintenance, logistics, worker guidance, data governance or the codification of knowledge. Intellimech performs an intermediary role by interpreting technological opportunities and connecting them to the operational language of companies. In orchestration terms, it converts scattered demand into a portfolio of possible challenges.

5.2 Shared knowledge creation and absorptive capacity building

The second mechanism is knowledge sharing. Intellimech presents itself as a network of knowledge, cooperation, constructive confrontation and exchange of ideas. The Knowledge Management Cluster provides a clear example (Intellimech, 2026). Public case histories indicate that Phoenix, SIAD Macchine Impianti and Tesmec participated in a cluster promoted by Intellimech together with ten other industrial companies. The initiative included a kick-off, shared workshops, individual meetings, periodic assessments and moments of comparison among firms.

The contents of the cluster are significant. Participating companies worked on issues such as management of non-conformities, introduction of dedicated software, standardization of project-related processes and

organizational activities. Phoenix emphasized the need to protect intangible assets and make knowledge available and resilient. SIAD described knowledge management as a way to protect, consolidate and make organizational information accessible to project actors. Tesmec connected knowledge management to operational continuity, shared memory and innovation.

This evidence suggests that Intellimech's orchestration role is not limited to technology transfer. It also concerns the design of organizational learning processes. By bringing companies together around knowledge management, the lab-consortium helps firms recognize common problems, compare practices and transform tacit issues into explicit improvement trajectories. This increases absorptive capacity: firms become better able to recognize, assimilate and apply new knowledge.

5.3 Pre-competitive experimentation and pilot projects

The third mechanism is experimentation. Intellimech promotes R&D projects and pre-competitive platforms, develops prototypes and participates in funded projects. The Watchman project illustrates this role. It aimed to develop a competence and experimentation hub for the application of AI techniques to machine vision in manufacturing, with demonstrators using deep learning to analyse images and recognize defects at the end of production processes. Intellimech acted as project leader in a partnership that included firms, a university and technology actors.

Other examples reinforce this pattern. The VIPS project concerned computer vision for quality control in small welds and included development of hardware, vision algorithms, image storage architecture and system integration. Shop4cf/Warning focused on augmented reality guidance for operators in agile production shopfloors, integrating HMI technologies, field data, enterprise data, CAD drawings and platform components to support assemblers and reduce non-conformities. Harmony, described as a regional project involving large firms, SMEs and research organizations, focuses on human-centric automation, immersive interfaces, ergonomics, human-robot collaboration and sustainable advanced manufacturing.

These examples show how Intellimech creates protected conditions for experimentation before full commercial adoption. Pre-competitive experimentation reduces the cost of learning for individual firms and creates reusable knowledge for the wider network. From an orchestration perspective, the lab designs arenas where firms can explore technologies with limited risk, observe demonstrations, participate in pilots and evaluate transferability to their own context.

5.4 Boundary spanning across firms, universities and institutions

The fourth mechanism is boundary spanning. Intellimech occupies a strategic position between companies, universities, institutions and innovation districts. Public information reports collaborations with universities including the University of Bergamo, Politecnico di Milano, University of Modena and Reggio Emilia, University of Calabria, University of Brescia and University of Pisa (Intellimech, 2026). It is also embedded in Kilometro Rosso, which is described as one of the main Italian innovation districts and as an open innovation environment with more than 80 resident partners and 2,500 employees and researchers.

This position allows Intellimech to act as a translator between different knowledge systems. Universities contribute scientific knowledge and research methods; firms contribute industrial problems and application contexts; institutional actors contribute legitimacy, funding channels and territorial development priorities; technology providers contribute tools and specialized competences. The orchestrator's role is to make these contributions combinable. It must reduce cognitive distance, create opportunities for interaction and ensure that projects remain relevant to firms while being technologically advanced.

Boundary spanning also supports the territorial dimension of innovation. Intellimech is not only a service provider for isolated firms; it contributes to a local and national industrial ecosystem by connecting companies from different sectors and by positioning them within broader innovation networks. This supports the idea that innovation labs can operate as ecosystem infrastructures rather than single-organization units.

5.5 Translation into firm-specific pilots and routines

The fifth mechanism concerns translation into firm-specific applications. AIxLAB provides a recent example. Intellimech describes AIxLAB as a laboratory designed to accompany companies through a structured path of AI exploration, learning and application. The Lovato Electric case highlights the value of combining training and practical experience to evaluate the real impact of AI in an industrial context and to compare with other companies on concrete application cases. The COES case similarly frames AIxLAB as a way to move from talking

about AI in the abstract to building something real and applicable to the firm, with areas such as internal process optimization, data analysis, decision support, logistics and embodied AI applied to production.

This mechanism is crucial because many innovation initiatives stop at awareness and inspiration. Intellimech's lab model appears to address this gap by connecting collective learning with firm-specific experimentation. Firms can participate in shared pathways and then identify applications that fit their own processes. In the Knowledge Management Cluster, this translation took the form of software implementation, standardization of project routines and management of non-conformities. In AIxLAB, it takes the form of AI applications embedded in business processes. In R&D projects, it takes the form of prototypes, demonstrators and pilots.

The contribution of the innovation lab is therefore not simply to produce technological artefacts, but to support the organizational translation required for adoption. This includes framing the problem, selecting the technology, designing the experiment, interpreting results, involving people and transforming learning into routines.

6 Discussion: A Model of Innovation Orchestration through Labs

The case suggests that innovation orchestration through a lab-consortium model can be understood as a cyclical process composed of five phases: sensing, framing, enabling, experimenting and embedding.

Sensing refers to the continuous identification of technological trends and member needs. Framing transforms broad trends into shared challenges and project opportunities. Enabling builds capabilities through training, workshops, mentoring, knowledge management tools and access to experts. Experimenting creates prototypes, demonstrators and pilots in pre-competitive or collaborative contexts. Embedding supports the translation of project learning into firm-specific routines, applications and strategic decisions.

This model extends the literature in three ways. First, it shows that an innovation lab can operate as a governance mechanism for open innovation. The lab does not only host creative sessions; it coordinates the interaction between firms, experts and institutions, helping companies select problems and organize experimentation. Second, it emphasizes the knowledge management dimension of innovation labs. Intellimech's activities around knowledge management, shared workshops, assessment and codification show that innovation orchestration involves the management of intangible assets, organizational memory and learning processes. Third, it highlights the economic and organizational sustainability of the membership-based model. The annual contribution of members supports a shared infrastructure that reduces the cost of technological exploration for individual firms while creating collective benefits.

The paper also shows that innovation labs can reconcile pre-competitive collaboration and firm-specific value creation. Companies may collaborate on generic technological themes, share learning and participate in clusters while still developing individual applications that support competitiveness. This is particularly relevant for SMEs, which may lack the resources to conduct continuous technology scouting or experimentation independently. By joining a lab-consortium, they gain access to a structured innovation environment and to the experience of other firms.

At the same time, orchestration is not neutral. It requires capabilities. The orchestrator must balance the interests of heterogeneous firms, manage confidentiality and intellectual property, select projects that are sufficiently shared but still valuable, and build trust among participants. The statute's attention to confidentiality and the protection of research results is consistent with this need (Intellimech, 2019). Without governance mechanisms, firms may be reluctant to share needs or participate in pre-competitive projects. Therefore, innovation orchestration combines technical competence, relational work and institutional design.

7 Implications

The study has implications for theory, managers and policy makers. For theory, it suggests that innovation labs should be studied not only as spaces for creativity but also as organizational platforms for knowledge orchestration. The concept of innovation orchestration helps explain how labs connect exploration, knowledge sharing and implementation. It also opens a dialogue between innovation lab research, open innovation and knowledge management.

For managers, the case highlights the value of participating in shared innovation infrastructures. Firms can benefit from technology scouting, structured learning pathways, mentoring, access to research expertise, project opportunities and comparison with peers. However, managers should not treat the lab as an external supplier only. The greatest value emerges when firms actively contribute needs, share experiences, involve internal teams and translate insights into internal projects.

For innovation lab managers, the case indicates the importance of designing a balanced portfolio of activities. Awareness events are useful, but they should be connected to workshops, assessment, pilots and implementation support. Similarly, technology projects should be complemented by knowledge management and capability-building activities. Labs that only develop prototypes risk producing isolated outputs; labs that also build routines and communities can generate more durable ecosystem effects.

For policy makers, the Intellimech case suggests that territorial innovation ecosystems can be strengthened through intermediary organizations that aggregate demand and mobilize firms around shared challenges. Public funding can support demonstration projects and experimentation hubs, but the sustainability of the ecosystem may depend on hybrid models that combine member contributions, funded projects, commissioned research and institutional partnerships.

8 Conclusions, Limitations and Future Research

This paper explored how innovation labs can orchestrate innovation in networks of firms and local ecosystems, using Intellimech as an illustrative case. The analysis suggests that Intellimech acts as a lab-consortium that enables companies to access knowledge, scouting, training, mentoring, open innovation opportunities and pilot projects. Its orchestration role unfolds through need sensing, knowledge sharing, pre-competitive experimentation, boundary spanning and translation into firm-specific routines.

The main conclusion is that the value of an innovation lab lies not only in technological experimentation, but in the organization of collective learning. Intellimech creates arenas where companies can define problems, compare experiences, access expertise and test solutions. This turns innovation from an isolated internal effort into a networked and knowledge-intensive process.

The paper has limitations. The analysis is based on public secondary evidence and does not include internal documents, interviews or direct observation. Therefore, the findings should be interpreted as an exploratory conceptualization rather than as a definitive evaluation. Future research should conduct interviews with Intellimech managers, technical staff and member firms, compare Intellimech with other innovation labs or digital innovation hubs, and examine how participation in the consortium affects firm-level innovation outcomes over time. Quantitative research could also investigate whether membership-based innovation labs increase absorptive capacity, project generation, technology adoption and business model innovation among participating firms.

Despite these limitations, the case provides a useful starting point for understanding innovation labs as orchestration infrastructures. In a context where firms must continuously navigate digital transformation and technological complexity, such infrastructures can help local ecosystems transform fragmented knowledge into coordinated innovation capacity.

9 Acknowledgements

The author acknowledges the publicly available information provided by Consorzio Intellimech and its member case histories. This draft is prepared for academic discussion and should be complemented, where possible, with primary data collection.

10 Ethics declaration

The study relies exclusively on publicly available documentary sources and does not involve human participants or personal data collection.

11 AI Declaration

Generative AI tools were used to support language editing. The author remains responsible for the accuracy, interpretation and final content of the paper.

References

Chesbrough, H.W. (2003), Open Innovation: The New Imperative for Creating and Profiting from Technology, Harvard Business School Press, Boston, MA.

Dhanaraj, C. and Parkhe, A. (2006), 'Orchestrating innovation networks', Academy of Management Review, Vol. 31 No. 3, pp. 659-669.

Enkel, E., Gassmann, O. and Chesbrough, H. (2009), 'Open R&D and open innovation: exploring the phenomenon', R&D Management, Vol. 39 No. 4, pp. 311-316.

Intellimech (2019), Statuto del Consorzio Intellimech in vigore, available at: https://www.intellimech.it/wp-content/uploads/2019/01/Statuto-Intellimech-in-vigore.pdf (accessed 1 June 2026).

Intellimech (2026), Home and consortium overview, available at: https://www.intellimech.it/ (accessed 1 June 2026).

Lewis, M. and Moultrie, J. (2005), 'The organizational innovation laboratory', Creativity and Innovation Management, Vol. 14 No. 1, pp. 73-83.

Magadley, W. and Birdi, K. (2009), 'Innovation labs: an examination into the use of physical spaces to enhance organizational creativity', Creativity and Innovation Management, Vol. 18 No. 4, pp. 315-325.

Nambisan, S. and Sawhney, M. (2011), 'Orchestration processes in network-centric innovation: Evidence from the field', Academy of Management Perspectives, Vol. 25 No. 3, pp. 40-57.

Nonaka, I. and Takeuchi, H. (1995), The Knowledge-Creating Company, Oxford University Press, New York, NY.

Santarsiero, F., Lerro, A., Carlucci, D., & Schiuma, G. (2022). Modelling and managing innovation lab as catalyst of digital transformation: theoretical and empirical evidence. Measuring Business Excellence, 26(1), 81-92.

Schiuma, G., and Santarsiero, F. (2024). Open Innovation Labs. *Cham, Switzerland: Springer.*

Spender, J.C. (1996), 'Making knowledge the basis of a dynamic theory of the firm', Strategic Management Journal, Vol. 17, Winter Special Issue, pp. 45-62.

West, J. and Bogers, M. (2014), 'Leveraging external sources of innovation: a review of research on open innovation', Journal of Product Innovation Management, Vol. 31 No. 4, pp. 814-831.

Governing AI in Academic Innovation Ecosystems: A Process Model from Italian University Regulations

Lorella Cannavacciuolo[1], Carmine Passavanti[2], Pierluigi Rippa[1], Giustina Secundo[3]

[1] Università degli Studi di Napoli Federico II, Naples, Italy
[2] Università Telematica Internazionale UNINETTUNO, Rome, Italy
[3] LUM Giuseppe Degennaro, Bari, Italy

Abstract

Artificial intelligence is rapidly transforming how innovation is organised within university-centred ecosystems. As AI-enabled practices diffuse across teaching, research, and collaborative activities, universities are required to redefine their governance arrangements in order to preserve transparency, accountability, and institutional legitimacy. These tensions are particularly visible in Academic Innovation Ecosystems (AIEs), where universities act as orchestrators coordinating interactions, infrastructures, and knowledge flows among diverse internal and external actors. While the literature increasingly acknowledges the relevance of responsible AI in higher education, we still lack a process-oriented understanding of how universities translate abstract principles into operational governance configurations embedded in ecosystem activities.

This study develops a conceptual model of AI governance in AIEs through qualitative document analysis grounded in abductive theorising. The empirical basis consists of official AI regulations and policy documents issued by Italian universities in response to emerging European regulatory frameworks. Out of 95 universities, 14 had adopted formal AI-related guidelines. These documents were analysed alongside institutional strategies and mission-related materials to capture how AI governance is articulated across education, research, and third-mission domains. The university is treated as the unit of analysis in its role as ecosystem orchestrator, with attention to how governance provisions shape the coordination of data and decision processes across missions.

Preliminary findings reveal a clear concentration of governance efforts around generative AI in teaching and research, whereas third-mission and administrative applications remain comparatively underdeveloped. In educational contexts, universities primarily focus on academic integrity and disclosure practices, whereas research-related provisions tend to emphasise authorship, intellectual property, and compliance with data protection norms. Across cases, regulatory references to European frameworks are present but uneven, and the operationalisation of monitoring mechanisms often remains generic rather than process-specific.

These patterns are theoretically significant because effective ecosystem orchestration depends on governance arrangements that are embedded in concrete practices rather than articulated solely as high-level principles. When AI governance remains reactive and unevenly specified across missions, the coordinating role of the university risks becoming fragmented. The proposed model advances research on innovation ecosystems by linking institutional conditions, infrastructural arrangements, and mission-specific governance choices, offering a foundation for more accountable and sustainable AI adoption within academic settings.

Keywords: Innovation Ecosystems; Academic Innovation Ecosystems; Digital Technologies; Artificial Intelligence; Conceptual Model.

Paper type: Academic Research Paper

1 Introduction

Digital technologies are transforming how innovation is organized, with artificial intelligence (AI) and platform infrastructures increasingly reshaping coordination and co-creation in complex ecosystems (Rippa and Secundo, 2019).

Within universities, these transformations are reflected in the emergence of Academic Innovation Ecosystems (AIEs), where the university orchestrates interdependent activities (knowledge creation, teaching, technology transfer, and academic entrepreneurship) so that knowledge is created, translated, and diffused through aligned actors, resources, data, and rules (Angrisani et al., 2023; Rippa & Secundo, 2019). We define an AIE as a university-centered socio-technical system in which digital infrastructures, institutional arrangements, and organizational processes interact to generate innovation and societal impact. AIE constitutes a distinctive research setting that integrates educational and research activities aligned with industrial demands, while also promoting engagement initiatives. Within an AIE, each actor plays a specific role connected to the three core missions of higher education institutions (namely, education, research, and third mission - Angrisani et al., 2023). Within this system, digital platforms and data governance provide the connective infrastructure that makes

participation feasible and supports modular growth across organizational and sectoral boundaries (Jarrahi et al., 2023). They provide the technical means to connect actors, standardize information flows, and reduce transaction costs. AI operates within these platforms as a functional and cognitive layer, augmenting human capabilities by enabling data-driven discovery, automation, and decision support (Jarrahi et al., 2023). In turn, Knowledge Management (KM) processes shape how data and insights produced by AI are transformed into actionable knowledge that sustains innovation within the ecosystem.

As academic and external partners move ideas across organizational boundaries, they rely on platform infrastructures that make data legible and interactions governable; within this infrastructure, AI works alongside human expertise to surface relevant knowledge when it is needed, to keep complex information connected to its context, and to sustain the everyday conversations through which opportunities are framed and refined (Secundo et al., 2024; Jarrahi et al., 2023). In practice, this means that searches become more targeted, documents and signals are linked through shared data structures, and routine exchanges are mediated in ways that preserve traceability and accountability, so that the movement from laboratory insight to collaborative development proceeds with greater clarity and speed (Primario et al., 2024; Rippa & Secundo, 2019). Put simply, platforms and governance compose the structural "rails" of the AIE, while AI provides the functional "engine" that raises the speed and quality of knowledge movement from labs to partners and markets (Secundo et al., 2024).

Existing reviews connect AI with collaboration and decision quality, yet they stop short of mapping these capabilities onto the concrete processes through which universities coordinate technology transfer, innovation programming, and academic entrepreneurship (e.g., Secundo et al., 2024; Rippa & Secundo, 2019).

Principles often recommended for AI in knowledge-intensive settings remain only partially institutionalized in the routines that govern university–industry collaboration, leaving a visible gap between guidance and practice (Jarrahi et al., 2023). This gap is amplified by governance choices, because arrangements such as hub-and-spoke consortia allocate data access, roles, and accountability in different ways, and this allocation shapes whether AI can be scaled responsibly in academic settings (Compagnucci et al., 2025; Daymond, 2022).

In response to this gap, we adopt a process lens: we examine AIE activities process by process and interpret them through distinct categories of AI, to describe systematically which mechanisms are activated and which benefits and risks emerge for different actors and levels; from this reading, we derive practical implications for orchestration and governance that align with the platform architectures and data infrastructures characteristic of university-centred collaboration.

In response to this gap, we examine the diffusion of AI within the university environment, whose landscape is characterized by three fundamental missions: teaching, research, and the third mission. These missions fuel the innovative ecosystem of AIE, which reconfigures education and research from an innovation-oriented perspective. Therefore, understanding the extent to which AI is transforming the ecosystem orchestrated by universities requires examining how universities are approaching this revolution.

Accordingly, this study aims to address the following research question: *how do universities configure AI governance within the three missions of universities?*

By addressing this question, the paper advances the understanding of AIEs in two main ways. *First*, it offers a process-oriented framework that connects ecosystem activities with AI capabilities, providing a structured way to analyze how knowledge is mobilized and transformed across academic and external boundaries. *Second*, it translates these insights into implications for university managers and policy makers, clarifying how investments in digital infrastructures, the design of organizational roles, and the implementation of accountability mechanisms can strengthen the functioning and legitimacy of AIEs

2 Theoretical Background

2.1 Academic Innovation Ecosystems (AIEs) in AI landscape

The concept of the AIE represents a novel and transformative reimagining of the integral function that universities serve in fostering and nurturing the intricate frameworks of innovation and entrepreneurial ecosystems that are pivotal in today's rapidly evolving economic landscape. The AIE is a physical environment in which university knowledge feeds the innovation and higher education coherently with the following value proposition: the creation of knowledge to be exploited outside the boundaries of the system (Angrisani et al., 2023). This model adopts a process-oriented perspective that emphasizes the dynamic and multifaceted interconnections that exist among knowledge, innovation, and entrepreneurship, which can be conceptualized as a complex structure comprising a web of relationships that involve multiple interacting actors across various domains. The university acts as an orchestrator connecting knowledge producers with business-oriented

laboratories and professional academies, aligning research, training, and application. The innovation ecosystem, orchestrated by university, enables knowledge to be captured, transformed and applied to create value.

The substantial value that is generated as a result of such innovations is systematically captured within the business ecosystem, thereby creating opportunities for exploitation and application in the broader external environment, which often includes market dynamics and competitive landscapes. Additionally, the transfer and dissemination of knowledge can be effectively executed through the advanced educational pathways that are situated within the AIE, which are thoughtfully provided by professional academies that are committed to academic excellence and industry relevance. It emerges that AIE concept not only facilitates the flow of knowledge but also significantly contributes to the overall advancement of both academic institutions and their industrial counterparts, thereby reinforcing the interconnectedness of education and economic growth.

In the AIE, the value generated by academic knowledge is systematically translated into external opportunities for innovation and entrepreneurship through a structured technology and knowledge transfer platform (Autio and Thomas, 2014). Such platform not only facilitates the exchange of ideas and resources but also fosters collaboration, thereby enhancing the overall capacity for innovation within the ecosystem, which is increasingly deemed indispensable in a knowledge-driven economy. The technology and knowledge transfer platform is administered within the academic environment, where a structured approach is adopted to ensure that the processes involved are both efficient and effective in meeting the overarching goals of innovation and development.

In this landscape, digital technologies, most notably AI, are reshaping innovation ecosystems by shifting coordination from primarily physical and geographic connections to digital and dynamic frameworks supported by platforms (Barile et al., 2025). Secundo et al. (2025) argue that AI acts as a catalyst for the universities to handle the growing complexity and volume of knowledge assets, optimize research collaboration, and drive responsible innovation within these ecosystems. The AI is transforming how the university's missions are performed because it affects the sharing of knowledge within classroom (first mission: education), the generation of novel knowledge and technology (second mission: research), and finally the knowledge transfer practices between University-Industry-Society (third mission).

2.2 AI as an Enabler of Knowledge Management

AI has become a transformative force in KM within university-based ecosystems, fundamentally reshaping how academic communities create, capture, share, and leverage information for research, education, and institutional development (Ruano-Borbalan, 2025). Both AI and KM are tightly linked to the nature of knowledge and learning; recent advances in AI provide new foundations for transforming KM in academic ecosystems (Rezaei, 2025). There are two complementary techno-organizational orientations: (1) KM, which manages knowledge within universities, and (2) AI, a branch of computing focused on developing systems that can emulate aspects of human knowledge and learning activities. We adopt a balanced, partnership perspective in which deep-learning AI augments rather than replaces human intellect, particularly for data-intensive, repetitive, and analytically complex tasks (e.g., forecasting or anomaly detection), while humans retain primacy over open-ended, ambiguous, or morally complex decisions (Jarrahi et al., 2023). This stance aligns with scholarship emphasizing AI-enabled augmentation in knowledge work, innovation processes, and decision-making systems where AI acts as a cognitive partner (Jarrahi et al., 2023). Hence, this perspective focuses on the partnership between humans and AI and on the unique capabilities of each for KM. The dynamic nature of academic knowledge demands more adaptive and intelligent KM solutions than traditional manual or database-driven systems. In this direction, AI-driven tools facilitate knowledge discovery, retrieval, and automation, thereby enhancing organizational knowledge utilization (Secundo et al., 2025). These technologies also contribute to the development of intelligent knowledge repositories, enabling organizations to bridge human expertise with AI-driven insights for superior decision-making and innovation (Ruano-Borbalan, 2025).

To better examine AI's role in supporting KM in universities, we adopt a process-oriented perspective (following Alavi and colleagues as discussed in Rezaei, 2025) and distinguish the following core KM processes: knowledge creation; storage and retrieval; sharing and collaboration.

Concerning knowledge creation, the generation of new knowledge often involves reconfiguring and recombining existing knowledge bases as much as developing ideas from scratch. AI-driven tools facilitate knowledge discovery by automating extraction, classification, and synthesis, turning raw academic data into actionable insights. Deep-learning AI supports predictive analytics for student performance and forecasting abandonment probabilities, enabling early interventions; administrative offices can also automate standard procedures that surround teaching processes. Machine learning algorithms identify patterns in large datasets, accelerate literature reviews, and enable meta-analyses that would otherwise be prohibitively time-consuming

for human researchers. NLP can parse academic articles, theses, and reports to extract concepts, findings, and methods, and automatically summarize key insights, expanding the frontier of inquiry.

Focusing on Knowledge Storage and Retrieval, creating and maintaining organizational memory requires strategies that extract knowledge, make it explicit, and record it systematically for future use via knowledge repositories. AI systems capture and curate this knowledge, powering intelligent repositories that bridge human expertise with AI-driven insights for superior decision-making and innovation. Adaptive learning systems leverage AI to tailor course content, assignments, and assessments based on each student's unique learning style, background, and progress. This not only enhances engagement and retention but also identifies and addresses learning gaps in real time.

Lastly, Knowledge collaboration is at the heart of university-based knowledge ecosystems, involving interdisciplinary research teams, external partners, and diverse communities of practice. AI revolutionizes collaboration through (Przegalinska et al, 2025):

- Real-time synchronization: Multi-user AI research platforms allow concurrent editing, instant updates, and version control, eliminating the confusion of traditional document sharing.
- Intelligent task distribution: AI allocates research and administrative tasks based on team members' expertise and availability, improving efficiency and knowledge sharing.
- Enhanced communication: AI integrates contextual information, prior discussions, and relevant literature into virtual meetings, facilitating more effective academic interactions.
- Such systems support complex group projects, streamline research workflows, and empower students and faculty to co-create knowledge dynamically.

Harnessing AI's benefits requires to face with critical questions about governance, ethics, and responsibility (Barile et al., 2025). Institutions should develop capabilities to ensure:

- *Data security*: AI systems managing sensitive research and student data must adhere to strict privacy and compliance standards, such as GDPR.
- *Bias mitigation*: Training data and algorithms should be regularly assessed to avoid reinforcing existing inequalities or under-representing certain groups.
- *Transparent decision-making*: AI's role in academic evaluation, resource allocation, and research assessment must be transparent and explainable to stakeholders.

This prompts universities to develop frameworks for responsible AI adoption that prioritize inclusivity, accountability, and social impact.

3 Method

This study is a conceptual contribution that develops a process model of AI-enabled AIEs through a rigorous qualitative document analysis combined with abductive theory building. Our logic of inquiry privileges mechanism-centered explanation over statistical inference and follows established guidance on documentary research by treating formal documents as empirical material that can be systematically collected and analyzed with a transparent audit trail (Secundo et al., 2024; Daymond, 2022; Bowen, 2009). We proceed in two coupled streams. In the first, we conduct a structured synthesis of the pertinent literature at the intersection of AIEs, digital platforms, and AI capabilities. We document the search strings, sources, and inclusion/exclusion criteria, and we iteratively consolidate constructs and definitions to ensure conceptual coverage and reproducibility (Secundo et al., 2024). In the second stream, we build an empirical documentary corpus on the governance and use of AI within Italian higher-education institutions. The starting point is the full population of Italian universities (state, non-state, and special-regime institutions). For each institution, we collect any formal AI regulation or guideline intended to "norm" AI use (at the time of data collection, 15 such regulations were available within a population of 95 institutions), and we complement these texts with institutional strategic plans and other official policy documents. We then extend the search to each university's official website to trace the concrete presence of AI-enabled tools or initiatives connected to the university's missions.

The unit of analysis is the university as an AIE orchestrator, operationalized through mission-specific instances of AI use. For each university, we code evidence separately for (i) education and capability building (first mission), (ii) research and knowledge valorization (second mission), and (iii) third mission activities (technology transfer, engagement, and academic entrepreneurship). Data extraction is performed through a shared, predefined coding matrix used by all authors to ensure consistency. For each mission-specific entry, the coder records: the AI use scenario; the AI category; any platforms or tools mentioned; the AIE processes involved; the relevant flows (data, knowledge, resources, decisions); expected or stated benefits and risks (classified as internal vs. external and direct vs. indirect); future development trajectories; document type; target audience; ethical principles invoked; implementation or monitoring mechanisms; and formal accountability

within the organizational structure. Each coded entry must be supported by a verbatim excerpt (with page number when available) or by a stable URL, and we record document provenance, publication date, and access date to maintain an audit trail (Roundy & Asllani, 2024).

To enhance reliability, two researchers independently code an initial subset of institutions to calibrate the codebook and resolve discrepancies, after which coding proceeds with periodic peer debriefs and documented revisions of the code definitions and decision rules (Daymond, 2022). Abductive iteration is operationalized through cycles in which emerging patterns in the coded matrix are compared against the evolving conceptual categories. Finally, we synthesize the coded evidence into within-university process traces that link institutional conditions, platform arrangements, and governance choices to observed or explicitly regulated changes in AIE processes and outcomes. We then apply cross-case pattern matching across universities to identify recurrent configurations that inform the proposed process model and its associated propositions.

4 Results and Discussion

Across the corpus, the strongest concentration concerns generative AI, especially in education and research. In the coded dataset, generative AI is explicitly addressed in 10 out of 14 universities for teaching related activities and for research related activities, while explicit attention to AI in third mission domains remains rare (1 out of 14) and only partially present in administrative functions (3 out of 14).

This distribution suggests that universities are reacting first to the most visible and rapidly diffusing practices, such as student use of generative tools in learning, and researcher use in writing and synthesis, while governance implications for technology transfer and academic entrepreneurship are still underdeveloped (Secundo et al., 2025; Rippa & Secundo, 2019).

In the teaching related provisions, governance concentrates on academic integrity and assessment legitimacy, typically via disclosure expectations and guidance on acceptable support versus substitution. In the research related provisions, the emphasis shifts towards authorship and compliance with intellectual property and data protection constraints, often treating AI outputs as aids that require verification and traceable attribution. Across both missions, benefits are typically framed as efficiency and quality gains, while risks are framed around privacy, bias, and erosion of trust in evaluation processes. (Jarrahi et al., 2023; Secundo et al., 2025).

Transversally, the documents emphasize high level principles and responsibility assignment, but they vary markedly in how far they translate principles into actionable mechanisms. References to major regulatory anchors are present but not uniform, with explicit mentions of the EU AI Act (EU AI Act, 2024) appearing in 6 out of 14 cases and GDPR (European Union, 2026) or privacy frameworks in 7 out of 14. This reinforces the idea that universities are constructing legitimacy through compliance narratives, yet the operationalisation of monitoring routines often remains generic. (Compagnucci et al., 2025; Secundo et al., 2025).

From an AIE perspective, this pattern is theoretically consequential because orchestration depends on the capacity to govern flows of data, knowledge, resources, and decisions across internal units and external partners and stakeholders (Angrisani et al., 2023). When AI governance is articulated mainly as a set of generic principles, and third mission and administrative domains remain weakly specified, the university's orchestration role risks becoming reactive and irregular across ecosystem processes.

To deepen the interpretation of these findings, we analyse AI governance through the lens of the KM cycle, distinguishing between knowledge creation, storage and retrieval, sharing and collaboration, and application (Alavi & Leidner, 2001; Grant, 1996). This perspective allows us to assess whether governance mechanisms align with the transformation of underlying knowledge processes within AIEs.

AI Governance in Knowledge Creation Processes

In both teaching and research domains, AI governance concentrates primarily on the knowledge creation phase. Universities frame generative AI as a support tool for idea generation. However, regulatory provisions do not primarily focus on redesigning collaborative knowledge production processes; rather, they concentrate on safeguarding the legitimacy and integrity of outputs.

In teaching-related provisions, governance centres on academic integrity and assessment legitimacy, typically through disclosure expectations and guidance on acceptable support versus substitution. In research-related provisions, the emphasis shifts towards authorship attribution, intellectual property, and compliance with data protection norms, often treating AI outputs as aids requiring human verification and traceable responsibility. Across both missions, benefits are typically framed as efficiency and quality gains, while risks are framed around privacy, bias, and erosion of trust in evaluation processes (Jarrahi et al., 2023; Ruano-Borbalan, 2025).

From a KM perspective, this indicates that AI governance is strongly output-oriented in the creation phase. While AI is acknowledged as augmenting cognitive work, governance mechanisms largely address accountability ex post rather than redesigning the socio-technical conditions of collaborative knowledge generation (Alavi & Leidner, 2001; Jarrahi et al., 2023). In this sense, the institutional response remains corrective rather than architectonic. Accordingly, when AI governance is concentrated in knowledge creation processes and framed around output legitimacy, universities reinforce reputational protection but limit the structural redesign of collaborative knowledge production systems (*Proposition 1*).

AI Governance and Organisational Knowledge Storage and Retrieval

Although less explicitly articulated, AI-related provisions also intersect with the storage and retrieval dimension of KM. References to data governance and compliance with GDPR suggest that universities are aware of AI's implications for organisational memory and information accessibility.

However, regulatory framing in this phase is predominantly compliance-driven. Explicit references to the EU AI Act appear in 6 out of 14 cases, while GDPR or privacy frameworks are mentioned in 7 out of 14. These references reinforce the role of legal anchors in constructing institutional legitimacy (Compagnucci et al., 2025; Secundo et al., 2025).

From the perspective of the knowledge-based view of the firm, organisational performance depends on the integration and coordination of distributed knowledge (Grant, 1996). In AIEs, digital infrastructures and data governance constitute the backbone through which knowledge becomes retrievable and scalable (Angrisani et al., 2023). We therefore propose that when AI governance is predominantly anchored in regulatory compliance rather than embedded in organisational knowledge infrastructures, institutional legitimacy strengthens but systemic integration of AI into knowledge architectures remains constrained (*Proposition 2*).

AI Governance in Knowledge Sharing and Boundary-Spanning Integration

The most pronounced asymmetry emerges in the sharing and collaboration phase of the KM cycle. Explicit governance attention to AI in third-mission activities appears in only 1 out of 14 universities. Administrative applications are mentioned in 3 out of 14 cases, typically in relation to internal efficiency rather than ecosystem coordination.

This imbalance is theoretically consequential: innovation ecosystems depend on distributed knowledge integration across organisational and sectoral boundaries (Autio & Thomas, 2014; West & Bogers, 2014). AIEs, in particular, rely on orchestrated flows between research units, teaching activities, and external stakeholders (Angrisani et al., 2023). When AI governance remains inward-looking and concentrated on classroom and publication contexts, the regulatory architecture fails to address how AI may reshape inter-organisational knowledge sharing and collaborative innovation.

Recent work on AI-enabled ecosystems emphasises the transformative role of AI in coordinating open innovation processes and balancing competition and collaboration (Barile et al., 2025; Primario et al., 2024). Yet our findings indicate that formal governance provisions lag behind this potential, especially in domains directly linked to entrepreneurial and third-mission activities. In KM terms, governance attention diminishes precisely where knowledge becomes more distributed and boundary-spanning. As knowledge processes become increasingly inter-organisational, the absence of formalised AI governance mechanisms heightens the risk of fragmented ecosystem coordination (*Proposition 3*).

AI Governance in Knowledge Application and Decision Integration

Across the corpus, documents emphasise high-level principles and responsibility assignment, but they vary in how far these principles are translated into actionable mechanisms. Monitoring routines, where mentioned, often remain generic, with references to committees or advisory bodies rather than clearly embedded procedural systems.

This pattern reinforces the idea that universities are constructing legitimacy through compliance narratives, while operationalisation remains uneven. From an ecosystem perspective, orchestration depends on the capacity to govern flows of data, knowledge, resources, and decisions across internal units and external partners (Daymond et al., 2023; Angrisani et al., 2023). When AI governance is articulated mainly as a set of generic principles and remains weakly specified in third-mission and administrative domains, the university's orchestrating role risks becoming reactive and irregular across ecosystem processes. So, the greater the alignment between AI governance mechanisms and the entire KM cycle (spanning creation, storage, sharing, and application) the stronger the university's capacity to orchestrate coherent knowledge flows within Academic Innovation Ecosystems (*Proposition 4*).

Taken together, the evidence reveals an asymmetric configuration of AI governance across the KM cycle. This asymmetry suggests that AI governance in AIEs is currently shaped more by reputational risk management and regulatory alignment than by a systemic redesign of knowledge processes. As AI increasingly becomes embedded in platform architectures and data infrastructures that sustain ecosystem coordination (Roundy &

Asllani, 2024; Secundo et al., 2025), the absence of process-integrated governance may generate fragmentation in how knowledge flows are structured and legitimised across the university's three missions.

To move from descriptive evidence to analytical structuring, we organise the findings into a configurational mapping of AI governance across KM cycle and university missions (Table 1).

Table 1: AI governance across KM cycle and university missions

	University Mission		
KM Process	***Education***	***Research***	***Third Mission & Administration***
Knowledge Creation	Strong regulatory attention to generative AI use in assignments, exams, and learning activities. Focus on disclosure, academic integrity, and legitimacy of outputs.	Strong attention to generative AI in drafting, data analysis, and synthesis. Emphasis on authorship, IP, traceability, and verification.	Minimal formalisation. AI rarely addressed in venture creation, technology transfer, or entrepreneurial knowledge generation.
Knowledge Storage & Retrieval	References to data protection, repository use, and privacy compliance. Limited strategic articulation of AI-enabled learning repositories.	Attention to research data governance, GDPR compliance, and data management practices. Limited discussion of AI-driven knowledge architecture redesign.	Sparse references to institutional memory or ecosystem-level knowledge repositories. Storage framed mainly through compliance logic.
Knowledge Sharing & Collaboration	AI framed as individual support tool rather than collaborative infrastructure. Limited attention to peer-to-peer AI-mediated collaboration.	Limited governance concerning AI in inter-institutional or interdisciplinary collaboration. Collaboration governance remains generic.	Very weak formalisation of AI in university–industry collaboration, open innovation, or entrepreneurial ecosystems. Boundary-spanning knowledge flows largely unregulated.
Knowledge Application & Decision Integration	Governance focuses on acceptable use rather than integration into decision processes.	AI framed as aid in research productivity, but limited embedding into structured decision architectures.	Few references to AI in administrative decision systems. Monitoring mechanisms remain high-level and principle-based.

5 Ethics Declaration

The authors declare that they have complied with all ethical requirements. No human participants or animals were involved in the study. No conflict of interest exist.

6 AI Declaration

The authors declare that ChatGPT was used exclusively for language editing and minor stylistic revisions. The tool was not used for conceptual development, data analysis, interpretation of results, or generation of substantive content. All intellectual contributions, arguments, and conclusions remain the sole responsibility of the authors.

References

Alavi, M., & Leidner, D. E. (2001). Knowledge management and knowledge management systems: Conceptual foundations and research issues1, 2. MIS quarterly, 25(1), 107-136.

Angrisani, M., Cannavacciuolo, L., & Rippa, P. (2023). Framing the main patterns of an academic innovation ecosystem. Evidence from a knowledge-intensive case study. International Journal of Entrepreneurial Behavior & Research, 29(11), 109-131.

Autio, E., & Thomas, L. (2014). Innovation ecosystems (pp. 204-288). The Oxford handbook of innovation management.
Barile, D., Secundo, G., & Del Vecchio, P. (2025). An artificial intelligence-based innovation ecosystem enabling open innovation and sustainable growth: evidence from a case study. Innovation, Organisation and Management, 1-23.
Compagnucci, L., Spigarelli, F., Perugini, F., Iacobucci, D., & Cobis, F. (2025). Does the hub and spoke model matter for university-industry engagement in innovation ecosystems?. The Journal of Technology Transfer, 1-32.
Daymond, J., Knight, E., Rumyantseva, M., & Maguire, S. (2023). Managing ecosystem emergence and evolution: Strategies for ecosystem architects. Strategic Management Journal, 44(4), O1-O27.
Grant, R. M. (1996). Toward a knowledge-based theory of the firm. Strategic management journal, 17(S2), 109-122.
Jarrahi, M. H., Askay, D., Eshraghi, A., & Smith, P. (2023). Artificial intelligence and knowledge management: A partnership between human and AI. Business Horizons, 66(1), 87-99.
Primario, S., Rippa, P., & Secundo, G. (2024). Peer innovation as an open innovation strategy for balancing competition and collaboration among technology start-ups in an innovation ecosystem. Journal of Innovation & Knowledge, 9(2), 100473.
Przegalinska, A., Triantoro, T., Kovbasiuk, A., Ciechanowski, L., Freeman, R. B., & Sowa, K. (2025). Collaborative AI in the workplace: Enhancing organizational performance through resource-based and task-technology fit perspectives. International Journal of Information Management, 81, 102853.
Rezaei, M. (2025). Artificial intelligence in knowledge management: Identifying and addressing the key implementation challenges. Technological Forecasting and Social Change, 217, 124183.
Rippa, P., & Secundo, G. (2019). Digital academic entrepreneurship: The potential of digital technologies on academic entrepreneurship. Technological Forecasting and Social Change, 146, 900-911.
Roundy, P. T., & Asllani, A. (2024). Understanding AI innovation contexts: a review and content analysis of artificial intelligence and entrepreneurial ecosystems research. Industrial Management & Data Systems, 124(7), 2333-2363.
Ruano-Borbalan, J. C. (2025). The transformative impact of artificial intelligence on higher education: A critical reflection on current trends and futures directions. International Journal of Chinese Education, 14(1), 2212585X251319364.
Secundo, G., Spilotro, C., Gast, J., & Corvello, V. (2025). The transformative power of artificial intelligence within innovation ecosystems: a review and a conceptual framework. Review of Managerial Science, 19(9), 2697-2728.
Secundo, Giustina, et al. "The transformative power of artificial intelligence within innovation ecosystems: a review and a conceptual framework." Review of Managerial Science 19.9 (2025): 2697-2728.
West, J., & Bogers, M. (2014). Leveraging external sources of innovation: A review of research on open innovation. Journal of product innovation management, 31(4), 814-831.

Controversial Advertising: Multicultural and Generational Perspectives on Knowledge and Cultural Meaning-Making

Gintarė Kriaučiūnaitė-Lazauskienė
Riga Technical University, Riga, Latvia

Abstract

Why do some advertisements inspire admiration and stick in people's minds across cultures, yet cause certain audiences to refuse to buy? This research investigates that puzzle by examining controversial advertising as a site of knowledge diversity. In this research, the perception of advertising is treated as a space in which culturally embedded unspoken knowledge and the collective construction of meaning intertwine with generational identity, thereby shaping how controversial messages are understood, interpreted, and responded to.

Twenty-two international students at Vilnius University of Applied Sciences (VIKO), Lithuania, the majority identifying as Generation Z, participated in a mixed-methods design combining focus group inquiry with a structured Likert-scale questionnaire. The sample was purposefully assembled to maximise cultural heterogeneity, with participants drawn from Morocco, Nigeria, Cameroon, Bangladesh, India, and Lithuania. The stimuli comprised five controversial advertisements drawn from Lithuanian and international contexts.

The main finding highlights what we term the "*paradox of advertising in shame-based cultures*". Students from collectivist and shame-oriented cultural backgrounds admitted that controversial advertisements attract attention and are easily memorable. However, they also emphasized that they would not want to purchase products or services from the advertised brands in their home countries. This resistance did not stem from personal indifference but was determined by communal knowledge systems that prioritise group reputation, family honour, and social conformity. Thus, the interpretation of advertising is never culturally neutral (de Mooij, 2019).

To explain this phenomenon, we present the *Cultural Knowledge Mediation Model* (CKMM). This five-stage system explains why traditional models of advertising effectiveness, such as the ELM, the AIDA hierarchy and the FCB Grid, may fail to adequately explain consumer reactions in collectivist and shame-based cultures. The model also generates three testable propositions (P1–P3) for future research.

For practitioners, the findings underscore that culturally intelligent, ethically responsible communication requires brands to engage with local knowledge ecosystems rather than assuming universal message reception. For knowledge management study, the study positions advertising reception as an under-examined knowledge-intensive phenomenon and challenges the prevailing treatment of Generation Z as a culturally homogeneous cohort.

Keywords: controversial advertising, cultural knowledge systems, generational meaning-making, knowledge diversity, social sustainability

Paper type: Academic Research Paper

1 Introduction

Consider the following scene: a group of university students from Morocco, Nigeria, India, and Lithuania sit together watching the same provocative advertisement. They all laugh; they all remember it afterwards. But ask them whether they would buy the product, and the room splits along cultural lines that no demographic variable such as age, education, digital fluency, can predict. This is not a failure of advertising effectiveness in any conventional sense but a collision between the message's intended knowledge transfer and the culturally embedded interpretive frameworks that audiences bring to the encounter.

This research examines controversial advertising communication: advertising deliberately designed to push the boundaries of social norms, raise sensitive issues, or challenge prevailing values through a knowledge management approach. This lens helps demonstrate that people do not merely react to controversial advertising. Their interpretation is based on what they already know, what they believe, and what they share with others. In this sense, the perception of controversial advertising is a process of knowledge formation. This tension forms the starting point of the research.

Organisations committed to socially sustainable communication need to know not just *whether* audiences attend to their messages, but *how* culturally diverse knowledge systems mediate the journey from attention to action – a question with direct implications for 'intelligent knowledge for sustainable organisations'.

The research reported here centres on 22 international students, predominantly Generation Z, at VIKO. Their diverse cultural backgrounds (covering North Africa, Sub-Saharan Africa, South Asia, and Eastern Europe)

presented a rare opportunity to observe how culturally distinct knowledge systems generate divergent responses to identical advertising stimuli.

2 Literature Review

2.1 Controversial advertising as a knowledge phenomenon

Controversial advertising is generally understood as commercial communication that provokes strong emotional reactions. It often employs prohibited elements such as taboo topics (religious elements, disgust, sexism) or socially sensitive contents (such as children, homeless pets) (Noland, 2021). Many studies have examined its impact on brand recognition, purchase intentions, and consumer attitudes (Machi et al., 2022; Hussain et al., 2023).

The process maps onto Nonaka and Takeuchi's (1995) SECI model of knowledge conversion. When a consumer encounters a controversial advertisement, the *Socialisation* phase activates shared tacit understandings about culturally embedded norms describing what is appropriate, what provokes shame, and what merits attention. Through Externalisation and Combination, these tacit reactions are articulated, shared and recombined with existing brand knowledge. Finally, *Internalisation* – the conversion of new explicit knowledge into personal implicit knowledge. This is a stage at which culturally embedded norms may block or facilitate brand adoption. Knowledge transfer barriers documented in organisational contexts (Argote and Ingram, 2000; Szulanski, 1996) apply with equal force to the consumer domain. Drawing on Grant's (1996) knowledge-based theory, we can understand culturally diverse audiences as possessing specialised knowledge that requires integration mechanisms which global advertising campaigns rarely provide. Table 1 (below) summarises prior empirical studies on Generation Z, advertising and multicultural knowledge practices.

Table 1: Selected empirical studies on Generation Z, advertising, and multicultural knowledge practices

Study	Context	Method / n	Ad Type	Key Finding (KM Relevance)
Li et al. (2026)	China	Qual. / n=14	Controversial sponsorship	Cultural conservatism rendered sexually suggestive content inappropriate; Gen Z transformed brand messages into memes as active sensemaking
Zulu et al. (2025)	South Africa	Quant. / n=504	Sustainable fashion	Cultural mindfulness positively influenced local brand attitudes; eco-label verification required to address trust deficits
Ali et al. (2026)	Indonesia	Quant. / n=306	Halal fashion	Religious identity more influential than product knowledge; collective norms dominated individual responses
Hu et al. (2026)	Unspec.	Exper. / n=509	CSR communication	Cultural ideology moderated CSR interpretation; intra-generational heterogeneity driven by ideology
Talesara (2025)	India	Mixed / n=115	Sustainable fashion	High environmental awareness but persistent attitude - behaviour gap; affordability as primary barrier
Fitriyana et al. (2025)	Unspec.	Qual. phenom.	Sustainability claims	Authentic narratives increased loyalty; inconsistent claims reduced purchase intent substantially
Pabian & Reformat (2025)	Poland	Qual. / n=70	Inclusive advertising	European Gen Z reported acceptance and empathy; responses distinct from Asian collectivist samples
Buckley et al. (2024)	Unspec.	Qual. / n=9	Femvertising	Identity alignment overrode inauthenticity; 'guilty displeasure' when criticising ads tied to gender identity
Sheng & Basha (2026)	China	Quant. / n=307	Short-video ads	Age/gender non-significant; content quality and algorithmic trust predicted purchase intent

Source: Structured literature overview (n=500 screened; 9 included); included studies span 2020–2026.

2.2 Cultural knowledge systems and the dynamics of interpretation

Culture is among the most powerful yet least explicitly theorised dimensions of knowledge in management scholarship. Hofstede et al.'s (2010) foundational work established that cultural values – such as individualism and collectivism, power distance, uncertainty avoidance, long-term orientation, and indulgence/restraint – systematically shape both organisational behaviour and communicative practice. Within advertising research, De Mooij (2019) has demonstrated that these dimensions influence the acceptability and effectiveness of provocative messaging across national contexts (Kunz et al., 2025).

For the present study, the concept of *shame culture* is particularly salient. In shame-oriented cultures, conduct is calibrated predominantly by perceived group judgement and the threat of social dishonour; in guilt-oriented cultures, internal moral evaluation plays a more prominent role (Benedict, 1946; Öztek et al., 2025). work on This distinction has been further illuminated by Markus and Kitayama's (1991) influential work, and more recently by Guo et al. (2024), emphasising *interdependent self-construal*: in collectivist societies, the self is constituted through social relationships rather than autonomous individuality, and behaviour is regulated through sensitivity to others' expectations rather than through internal guilt. Hall's (1976) framework adds a further lens: shame cultures typically operate as high-context environments, where meaning is embedded in social relationships and situational cues rather than in explicit verbal content. Despite its explanatory potential, the shame–guilt dimension has received limited systematic attention in advertising research, a gap this study directly addresses.

What makes shame culture so consequential for advertising knowledge transfer is the nature of the knowledge involved. We are dealing not with mere individual preference but with what Polanyi (1966) calls *tacit knowledge* – knowledge that is deeply personal, resistant to explicit articulation, and acquired through immersion in a community of practice. For individuals from shame-culture backgrounds, such advertisements are not merely distasteful; they activate community-level knowledge systems bound up with honour, reputation and social presentation, making the interpretive act collective even when the viewer sits alone.

2.3 Generation Z: Cohort unity, cultural diversity

Generation Z (in this researched aimed from 1999-2003) has attracted significant attention from researchers as a consumer group, recognized as having grown up in the digital age characterized by high levels of media literacy and an interest in social and ethical issues (Dwivedula, 2025). Several researchers argue (Seemiller and Grace, 2019; Dolot, 2018) that this generation demonstrates greater openness to diversity, social issues, and non-normative content than previous generations.

Yet a troubling simplification pervades much of this literature: the treatment of Generation Z as a globally homogeneous group. This assumption obscures the profound cultural and geographical diversity within the cohort. As Dolot (2018) rightly cautions, generational characteristics are not culturally neutral. Rather, they reflect the specific social, educational, and familial environments in which individuals develop. For international students, who must navigate multiple cultural frameworks simultaneously, the tension between generational identity and cultural heritage becomes particularly acute.

2.4 Thematic synthesis

Four themes emerge from the nine included studies that reinforce and extend the theoretical claims above. *First*, cultural knowledge systems consistently override generational cohort characteristics as the primary determinant of advertising interpretation. Li et al. (2026) find that Chinese Gen Z rejected sexually suggestive content despite generational openness, while Ali and Rafdinal (2026) show that Indonesian religious community identity outweighed individual product knowledge as a pattern structurally analogous to the shame-culture dynamics central to this study, and consistent with Bagozzi, Verbeke and Gavino's (2003) finding that culture moderates the self-regulation of shame. *Second*, the evidence substantiates the attitude –behaviour dissociation at the heart of what we term the shame-culture paradox. Talesara (2025) demonstrates cases where stated values failed to predict purchasing behaviour. *Third*, Generation Z emerges as a knowledge-active cohort. Li et al.'s (2026) meme-making esports fans and Buckley et al.'s (2024) "guilty displeasure" among femvertising viewers both illustrate collective sensemaking. *Fourth*, the sustainability implications of culturally intelligent communication are empirically robust: Fitriyana et al. (2025) and Zulu et al. (2025) demonstrate that authentic, culturally attentive communication generates measurable gains in brand trust and loyalty.

2.5 Conceptual framework: The Cultural Knowledge Mediation Model

The theoretical strands reviewed above converge on a problem that existing models leave unresolved: why do conventional advertising effectiveness frameworks – such as the Elaboration Likelihood Model (Petty and Cacioppo, 1986), the AIDA hierarchy, the FCB Grid (Vaughn, 1980) – fail to predict advertising responses in collectivist, shame-oriented cultural contexts? These models implicitly assume culturally neutral, individual-level processing; the empirical evidence reviewed here suggests that this assumption breaks down precisely where culturally embedded knowledge systems intervene.

To address this gap, we propose the *Cultural Knowledge Mediation Model (CKMM)*, an integrative five-stage framework synthesising knowledge management theory (Nonaka and Takeuchi, 1995; Tsoukas, 1996; Grant, 1996), cross-cultural psychology (Hofstede et al., 2010; Benedict, 1946; Hall, 1976; Markus and Kitayama, 1991).

The model operates as follows (Figure 1). *Stage 1:* A controversial advertising stimulus enters the field. *Stage 2:* Before cognitive processing occurs, the stimulus passes through a *cultural knowledge* filter shaped by the audience's position on Hofstede's individualism–collectivism dimension, Benedict's shame–guilt continuum, and Hall's high–low context communication spectrum. *Stage 3:* The filtered stimulus enters the *knowledge transfer process*, theorised through Nonaka and Takeuchi's SECI model. In shame-culture contexts, the Socialisation phase activates tacit peer and elder norms, creating barriers to Internalisation of culturally misaligned brand associations. *Stage 4:* The model produces two parallel but potentially divergent outputs: cognitive responses and behavioural intent. This division is the CKMM's central departure from conventional models. *Stage 5:* All pathways converge on social sustainability outcomes: responsible brand communication, brand trust, and long-term stakeholder legitimacy.

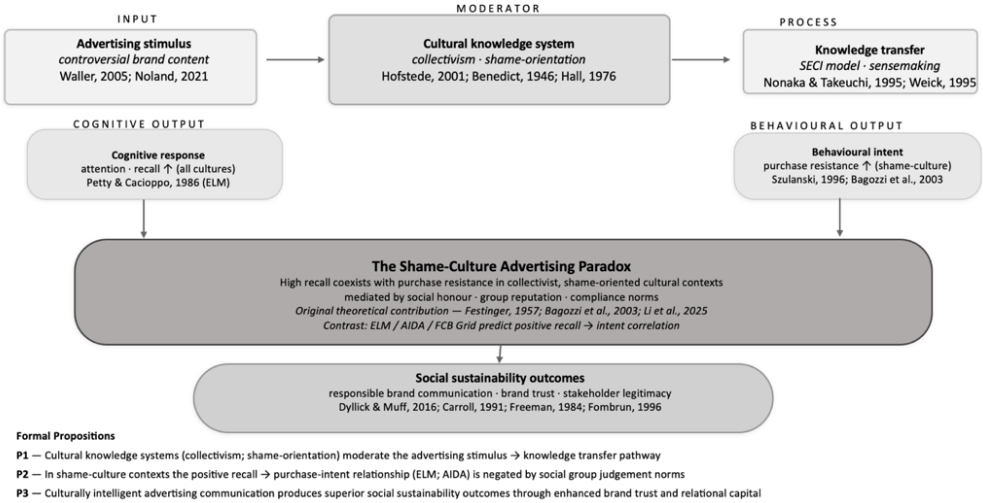


Figure 1: Cultural Knowledge Mediation Model (CKMM).
Source: Author's own conceptualisation (2026). P1–P3 = formal propositions derived from the framework.

The CKMM's central analytical contribution is the formalisation of the shame-culture advertising paradox. The paradox emerges at the bifurcation between cognitive and behavioural outputs: high recall coexists with strong purchase resistance because the knowledge transfer process is blocked by community-level knowledge systems related to honour and social compliance. This constitutes a cognitive dissonance dynamic in Festinger's (1957) terms – the individual's private receptivity is in tension with the publicly required performance of non-engagement. Here, three formal propositions follow.

P1: The more collectivist and shame-oriented an audience's cultural knowledge system, the larger the gap between advertising recall and purchase intention for controversial stimuli. P2: The recall–intention gap in P1 widens in intergenerational or public viewing contexts and narrows in private or peer-only contexts. P3: Longer exposure to individualist host cultures (acculturation) progressively attenuates the shame-culture effect.

3 Methodology

The study employed a mixed-methods design combining qualitative focus group inquiry with a supplementary quantitative questionnaire (Creswell and Creswell, 2018). This combination enables rich exploration of meaning-making processes while providing systematic attitudinal data amenable to comparison across cultural groups.

3.1 Participants

Twenty-two international students enrolled at VIKO. Participants were purposefully selected to maximise cultural diversity (voluntary), with representation from Morocco, Nigeria, Cameroon, Bangladesh, India, and Lithuania (as a comparative group). The majority identified as Generation Z (born 2001–2004, within the generational range defined in Section 2.3).

3.2 Focus group

The two focus group sessions, lasting approximately 75–90 minutes, were conducted at VIKO. Participants were shown five controversial advertisements drawn from Lithuanian and international contexts, encompassing sexually suggestive imagery, sexual wordplay in food advertising, destination marketing, and personal care campaigns. Following Krueger and Casey (2015), a semi-structured discussion guide elicited responses concerning emotional reactions, cultural values, ethical judgements, and purchase intentions. The session was in English, audio-recorded and transcribed verbatim.

3.3 Questionnaire

A short, structured questionnaire was administered after the focus group. Eight statements were measured on a five-point Likert scale (1 = Strongly disagree; 5 = Strongly agree), covering attentional capture, memorability, cultural influence on perception, generational openness, brand harm, purchase intention, and the need for cultural adaptation.

3.4 Analysis

Focus group transcripts were analysed using Braun and Clarke's (2023) reflexive thematic analysis. Initial codes were generated inductively and iteratively organised into candidate themes, with priority given to patterns relating to cultural knowledge systems, sensemaking processes, and the relationship between advertising reception and cultural identity. Questionnaire data were summarised using descriptive statistics. Integration followed a convergent parallel design (Creswell and Creswell, 2018), in which qualitative findings provide interpretive depth for quantitative patterns.

4 Findings

4.1 Attentional impact and knowledge activation

Across all cultural backgrounds, participants confirmed without exception that controversial advertising commands attention and generates enduring recall. Questionnaire data reflected this: "Controversial advertising attracts my attention" scored a mean of 4.1, and "I remember controversial ads more than regular ads" a mean of 3.9. In the focus group, participants described controversy as something that compels engagement almost involuntarily – a stimulus that activates epistemic curiosity and social debate rather than passive viewing. This foundation supports the SECI logic: schema-disrupting stimuli push audiences into the Socialisation and Externalisation phases, producing deeper processing and more durable retention than schema-confirming information.

4.2 The shame-culture advertising paradox

The most distinctive finding was a consistent pattern among participants from North African, Sub-Saharan African, and South Asian backgrounds. These students acknowledged the attentional power of controversial advertisements while simultaneously expressing firm resistance to purchasing the featured brands in their home-country contexts. Crucially, this resistance was not about personal distaste; it was about community.

A participant from Morocco described experiencing no personal discomfort with certain advertisements but explained that encountering such content alongside parents or community elders would necessitate either physical withdrawal or a deliberate performance of non-engagement. A participant from Cameroon used the phrase "you would be ashamed" to capture the contextual discomfort of encountering sexually suggestive advertising in intergenerational settings. Participants from Bangladesh and India drew explicit connections to family and caste structures as knowledge systems that define acceptable brand associations.

What we see here is the shame-culture advertising paradox: high memorability coexisting with active purchase resistance rooted in *interdependent self-construal*.

4.3 Contextual code-switching as knowledge practice

Participants exhibited considerable sophistication in modulating their responses according to social context. Several described markedly different reactions to the same advertisement depending on whether it was encountered alone, with peers, or in an intergenerational family setting. The statement "*I would act like I didn't see it*", offered by multiple participants when describing intergenerational viewing captures a widely shared strategy of contextual code-switching.

This behaviour reveals that advertising reception is not a fixed individual response but a socially situated knowledge practice. The meaning attributed to controversial content is dynamically reconstructed depending on the knowledge context in which it is encountered – precisely the kind of contextual, retrospective interpretation that Weick's (1995) sensemaking framework predicts. Consistent with Hall (1976), participants from high-context cultures derived meaning not from the advertisement's explicit content alone but from the social situation surrounding its reception.

4.4 Intra-generational diversity

While participants broadly agreed that their generation is more open to controversial advertising than older cohorts (mean = 3.7), the qualitative evidence told a more nuanced story. Participants from collectivist backgrounds displayed markedly more cautious and contextually conditional acceptance than those from more individualist settings. Generational membership supplies a shared communicative register; cultural knowledge systems supply the interpretive filters through which it operates. Generation Z's knowledge-active characteristics do not override cultural heritage; they interact with it in complex, sometimes contradictory, ways (see Table 2).

Table 2: Summary of questionnaire attitudinal items (n = 22)

Statement	Mean	SD
Controversial advertising attracts my attention	4.1	0.82
I remember controversial ads more than regular ads	3.9	0.94
Some controversial ads make me feel uncomfortable	3.8	1.05
Cultural background influences how I perceive controversial ads	4.3	0.71
My generation is more open to controversial advertising	3.7	0.98
Controversial advertising can harm a brand's image	3.6	1.02
I would stop buying from a brand after a controversial ad	3.1	1.14
Brands should adapt controversial ads for different cultures	4.4	0.68

Source: Author's own research (2026). Scale: 1 = Strongly disagree; 5 = Strongly agree.

The two highest-scoring items concerned cultural influence on perception (4.3) and the need for cultural adaptation (4.4) strengthening the qualitative evidence for the centrality of cultural knowledge systems. The comparatively lower mean for purchase discontinuation (3.1) is telling: participants recognise cultural discomfort without necessarily translating it into universal brand avoidance, a pattern consistent with the shame-culture paradox's distinction between private attitude and public behaviour. Remarkably, this item measured general boycott intention, while the paradox concerns context-specific purchase resistance in home-country and intergenerational settings, which the qualitative data capture more directly.

5 Discussion

The findings carry several implications for knowledge management theory and for the practice of global brand communication.

Controversial advertising as knowledge-intensive phenomenon. The study establishes that the effectiveness of controversial advertising cannot be understood without reference to the tacit cultural knowledge (Polanyi, 1966) that audiences bring to the interpretive encounter. The SECI model illuminates the specific failure point: in shame-culture contexts, the Socialisation phase activates community norms that block Internalisation – the stage at which brand associations would normally become embedded in individual knowledge repertoires. Organisations deploying controversial communication without mapping the tacit knowledge landscapes of their audiences risk producing what Szulanski (1996) terms 'knowledge stickiness': the message arrives but cannot be absorbed because of causal ambiguity about the brand's cultural intent and motivational barriers rooted in collective identity.

The shame-culture paradox as a structural challenge to advertising theory. Conventional advertising effectiveness models assume a positive relationship between affective engagement and purchase intention. Our findings demonstrate that this relationship is culturally conditional: in shame-culture contexts, high memorability coexists with strong purchase resistance when the advertising stimulus activates knowledge systems related to honour and social compliance. The evidence in Section 2.4 reveals this as a cross-contextual pattern: Talesara (2025) demonstrates analogous attitude–behaviour dissociations driven by economic constraints, suggesting that the paradox reflects a broader structural feature of Gen Z consumer behaviour that takes a distinctively cultural-normative form in collectivist settings.

The limits of generational framing. Much of the literature treats Generation Z as a globally homogeneous cohort; our findings complicate this picture considerably. Cultural knowledge systems, particularly those associated with collectivist, shame-oriented structures and *interdependent self-construal* (Markus and Kitayama, 1991), exert a persistent influence on advertising responses even among digitally literate, globally connected young adults. The contextual code-switching shown here, where participants reacted very differently to the same advertisement depending on the social situation as prove that the meaning of an advertisement is not fixed. Rather, it is co-constructed by the message, the audience's beliefs and knowledge, and the context in which it is encountered (Khan & Kumar, 2026; Anker, 2024).

6 Conclusions

This study has examined controversial advertising through the lens of knowledge management, revealing how culturally embedded knowledge systems and generational dynamics interact to produce diverse and sometimes paradoxical patterns of response. Its central contribution is the identification and formalisation of the *shame-culture advertising paradox*: a consistent pattern in which high attentional and memorial impact coexists with purchase resistance among individuals from collectivist, shame-oriented backgrounds. It becomes intelligible only when advertising reception is understood as a culturally mediated knowledge process, formalised here through the Cultural Knowledge Mediation Model (CKMM) and its three testable propositions.

The structured literature overview (nine included studies) provides independent cross-cultural support for the paradox and shows that analogous value–behaviour dissociations recur in economically and ideologically diverse Gen Z samples.

For practitioners, the message is clear: culturally intelligent communication is not a nicety but a necessity for brands seeking both commercial effectiveness and social legitimacy. For scholars, the study positions advertising reception as a knowledge-intensive phenomenon deserving sustained attention within KM allowance and challenges the assumption that Generation Z is a culturally unified audience.

Several limitations suggest productive future directions. The sample, while culturally diverse, is small (n = 22) and drawn from a single institutional setting, and the single, large focus group format may have amplified dominance and social-desirability effects. Replication with larger, more geographically varied samples is essential. Future work should examine how the shame-culture paradox operates across other controversial advertising categories and generational cohorts. The acculturation variable merits particular attention: Berry's (1997) framework for immigration and adaptation suggests that time spent outside one's home culture may progressively moderate the shame-culture dynamic, a hypothesis that longitudinal designs are well suited to test.

7 Acknowledgements

The author thanks the international students at Vilnius Kolegija who participated in this research.

8 Ethics Declaration

Ethical clearance was obtained at the institutional level. All participants provided written informed consent, were assured of full anonymity, and were informed of their right to withdraw without consequence. Audio recordings and transcripts are stored on password-protected devices and will be retained only as needed for academic research purposes, in compliance with applicable data protection regulations.

9 AI Declaration

Artificial intelligence writing assistance tools Grammarly and DeepL were used for language refinement. The research design, data collection, empirical analysis, and intellectual arguments are entirely the work of the author.

References

Ali, D. S. F., & Rafdinal, W. (2026). Fashioning hyperpersonal relationship in halal fashion brand: Halal brand loyalty and social information processing. *Journal of Islamic Marketing*, *17*(3), 1058–1079. https://doi.org/10.1108/JIMA-06-2024-0234

Anker, T. B. (2024). Meaningful choice: Existential consumer theory. *Marketing Theory*, *24*(4), 591–609. https://doi.org/10.1177/14705931231207317

Argote, L., & Ingram, P. (2000). Knowledge transfer: A basis for competitive advantage in firms. *Organizational Behavior and Human Decision Processes*, *82*(1), 150–169.doi: 10.1006/obhd.2000.2893

Bagozzi, R. P., Verbeke, W., & Gavino, J. C., Jr. (2003). Culture moderates the self-regulation of shame and its effects on performance: The case of salespersons in the Netherlands and the Philippines. *Journal of Applied Psychology*, *88*(2), 219–233. https://doi.org/10.1037/0021-9010.88.2.219

Benedict, R. (1946). *The chrysanthemum and the sword: Patterns of Japanese culture*. Houghton Mifflin.

Berry, J. W. (1997). Immigration, acculturation, and adaptation. *Applied Psychology*, *46*(1), 5–34.

Braun, V., & Clarke, V. (2023). Is thematic analysis used well in health psychology? A critical review of published research, with recommendations for quality practice and reporting. *Health Psychology Review*, *17*(4), 695–718. https://doi.org/10.1080/17437199.2022.2161594

Buckley, A. C., Yannopoulou, N., Gorton, M., & Lie, S. (2024). Guilty displeasures? How Gen-Z women perceive (in)authentic femvertising messages. *Journal of Current Issues & Research in Advertising*, *45*(4), 388–401. https://doi.org/10.1080/10641734.2024.2305753

Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications.

de Mooij, M. (2019). *Consumer behavior and culture: Consequences for global marketing and advertising* (3rd ed.). SAGE Publications.

Dolot, A. (2018). The characteristics of Generation Z. *e-Mentor*, *74*(2), 44–50.

Dwivedula, R. (2025). Why is Generation Z motivated at work? A qualitative exploration. *Global Business and Organizational Excellence*, *44*(3), 38–49. https://doi.org/10.1002/joe.22273

Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford University Press.

Fitriyana, N. A., Suryadi, N., & Hapsari, R. D. V. (2025). The role of social media and online reviews in building brand loyalty through trust. *Jurnal Ilmiah Manajemen Kesatuan*, *13*(5), 3324–3336. https://doi.org/10.37641/jimkes.v13i5.3602

Grant, R. M. (1996). Toward a knowledge-based theory of the firm. *Strategic Management Journal*, *17*(S2), 109–122.

Guo, T., Cai, C., Xu, Y., Huang, M., & Ni, Y. (2024). Self-esteem and college students' online impulse buying: The independent and interactive moderating role of interdependent self-construal and interpersonal relationships. *Psychology Research and Behavior Management*, *17*, 4447–4461. https://doi.org/10.2147/PRBM.S494636

Hall, E. T. (1976). *Beyond culture*. Anchor Press.

Hofstede, G., Hofstede, G. J., & Minkov, M. (2010). *Cultures and organizations: Software of the mind* (3rd ed.). McGraw-Hill.

Hu, Y., Ma, Y., & Mu, D. (2026). How Gen Z makes sense of CSR communication: The role of perceived altruism and cultural ideology in responses to internal and external CSR. *Young Consumers*. Advance online publication. https://doi.org/10.1108/YC-04-2025-2515

Hussain, A., Hooi Ting, D., Zaib Abbasi, A., & Rehman, U. (2023). Integrating the S-O-R model to examine purchase intention based on Instagram sponsored advertising. *Journal of Promotion Management*, *29*(1), 77–105. https://doi.org/10.1080/10496491.2022.2108185

Khan, S., & Kumar, P. (2026). Meaning in context: Interpreting brand communication through cultural frames and generational identity. *Studies in Science of Science*, *44*(1), 427–442.

Krueger, R. A., & Casey, M. A. (2015). *Focus groups: A practical guide for applied research* (5th ed.). SAGE Publications.

Kunz, S., Florack, A., Leri, I., Schmied, A., & Alves, H. (2025). The challenge to be different in influencer marketing: Why and when negative meanings of influencers transfer to brands. *International Journal of Advertising*, *44*(7), 1341–1365. https://doi.org/10.1080/02650487.2024.2430861

Li, S., Wong, D., Meng-Lewis, Y. (2026), "Negotiating sexuality, identity and marketplace culture: a case study on Gen Z's interpretations of Durex's esports sponsorship in China". *International Journal of Sports Marketing and Sponsorship*, Vol. 27 No. 2 pp. 431–449, doi: https://doi.org/10.1108/IJSMS-04-2025-0178

Machi, L., Nemavhidi, P., Chuchu, T., Nyagadza, B., & de Villiers, M. V. (2022). Exploring the impact of brand awareness, brand loyalty and brand attitude on purchase intention in online shopping. *International Journal of Research in Business and Social Science*, *11*(5), 176–187. https://doi.org/10.20525/IJRBS.V11I5.1841

Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, *98*(2), 224–253. https://doi.org/10.1037/0033-295X.98.2.224

Noland, C. R. (2021). Positive or negative vibes? Does mood affect consumer response to controversial advertising? *Journal of Marketing Communications*, *27*(8), 897–912. https://doi.org/10.1080/13527266.2020.1795911

Nonaka, I., & Takeuchi, H. (1995). *The knowledge-creating company: How Japanese companies create the dynamics of innovation*. Oxford University Press.

Öztek, M. Y., Sağlam, M. H., & Türk, E. (2025). Longitudinal study of perceived brand globalness: The dynamic effects of ethnocentrism and purchase intentions from 2021 to 2024. *Sustainability*, *17*(15), Article 7132. https://doi.org/10.3390/su17157132

Pabian, B., & Reformat, B. (2025). Perception of inclusive advertising by representatives of Generation Z. *Zeszyty Naukowe Politechniki Śląskiej. Organizacja i Zarządzanie, 218*, 527–535. https://doi.org/10.29119/1641-3466.2025.218.29

Seemiller, C., & Grace, M. (2019). Generation Z: A century in the making. Routledge.

Petty, R. E., & Cacioppo, J. T. (1986). *Communication and persuasion: Central and peripheral routes to attitude change*. Springer.

Polanyi, M. (1966). *The tacit dimension*. Doubleday.

Sheng, M., & Basha, N. K. (2026). The impact of short video advertisements on online purchasing intentions among Chinese Gen Z consumers. *Malaysian Journal of Social Sciences and Humanities (MJSSH)*, *11*(1), Article e003694. https://doi.org/10.47405/mjssh.v11i1.3694

Szulanski, G. (1996). Exploring internal stickiness: Impediments to the transfer of best practice within the firm. *Strategic Management Journal*, *17*(S2), 27–43.

Talesara, A. (2025). *To explore and analyse the attitude, behaviour, and purchasing decisions of Generation Z (Gen Z) consumers towards sustainable fashion*. International Journal of Scientific Research in Engineering and Management, 9(4), 1–9. https://doi.org/10.55041/IJSREM45479

Tsoukas, H. (1996). The firm as a distributed knowledge system: A constructionist approach. *Strategic Management Journal*, *17*(S2), 11–25. https://doi.org/10.1002/smj.4250171104

Vaughn, R. (1980). How advertising works: A planning model. *Journal of Advertising Research*, *20*(5), 27–33.

Weick, K. E. (1995). *Sensemaking in organizations*. SAGE Publications.

Zulu, M., Muposhi, A., Matselale, N., & Parbhoo, N. (2025). Juxtaposing Generation Z consumers' consumption behaviour of local versus global sustainable fashion brands. *Young Consumers*, *26*(7), 229–253. https://doi.org/10.1108/YC-12-2024-2357

INDEX – ASIA CHAPTER

1349 *Pao-Jung Chao, Szu-Ying Lai, Chih-Chieh Chang*
Exploring User Intentions and Behavior Towards Generative AI Adoption in Conversational Commerce

1357 *Tzu-Hsuan Wang, Yi-Shan Chen, Chih-Chieh Chang*
Exploring Federated Learning Adoption in Competitive and Collaborative Interorganizational Environments: A TOE-S Perspective

1366 *Chiung-Chuan Lo, Ya-Ping Lin*
Framework Development of an ESG Course Series under System Dynamics: A Comparison of TKS and KSA Perspectives

1375 *Ju-We Chen, Chi-Yuan Lin, Tsu-Erh Kuo*
A System Dynamics Analysis of AI-Empowers Smart Aquaculture Technologies and Their ESG Performance Impacts

1384 *Chiung-Chuan Lo*
A System Dynamics Analysis of AI-Driven Competency Development in Business Schools

1394 *Hsuan-Hao Lo, Sirirat Sae Lim*
Evolution of Knowledge Landscapes in AI-KM: A Longitudinal Analysis of Divergence between Term Diversity and Community Attention (2015-2024)

1403 *MeiYi Winnie Song, Ju-We Chen*
Embedding ESG in Strategic Management Course: A System Dynamics Approach to Experiential Learning

1413 *MeiYi Winnie Song , Kuan-Lun Chen*
A System Dynamics Analysis of Career Agency Formation among Gen Z Interns: Unpacking the Contextual Mechanisms of Overseas Work-Integrated Learning

1422 *Ju-We Chen, Shin-Ping Su*
AI-Empowers Smart Aquaculture: A System Dynamics-Integrated Decision-Making Architecture

1432 *Yi-Chen Cheng, Ching-Yeh Chu*
Inclusive Governance of Sport Centers: A System Dynamics Perspective

1442 *Chia-Fen Lin, Yin-Feng Chen, Rou-Ya Lin*
Towards Smart and Sustainable Preschools: System Dynamics Modeling of AI-empowered Educational Governance under ESG Principles

1452 *Bin-Wei Hsu, Hsiu-Yi Lin*
How Does Knowledge Externalization Quality Shape the ESG–Financial Performance Nexus: An Explainable AI Analysis under Mandatory Disclosure Regime

1462 *Sophie Kung, Nicole Lin, Joy Hong, Peter Lee, Yaping Tsao*
Deepfake, Deep Scars-Protecting Children's Rights in the Age of Artificial Intelligence

1469 *Kai-En Chang, Yo-Yo Chen, Yun-Hsuan Chu*
Sexual Violence Towards Female Domestic Migrant Workers

1478 *Huang, Yu-En, Wu, Mei-Ling, Cheng, Hsiang-Wei*
Modern Dystopia: Social Media and Digital Fragility

1486 *Jimmy Mao, Leticia Hsu, Sophie Yang*
When Wetlands Disappear: What Human Development Means for Nature?

1495 *David Coles, Tzu-Ting Hsu, Min-Ren Yan*
Montessori Education as a Human-Centered Foundation for AI Competency Development

1504 *Duong Minh Tran, Wen-Min Lu*
The Relationship between ESG and Firm Efficiency with the Moderating Role of Digital Transformation and Knowledge Capability

1512 *Huang, Yan-Ping*
Developing an Intelligent Organizational Carbon Inventory Assistant Integrating LangChain and Large Language Models: An Empirical Study

1521 *Huang, Yan-Ping Huang, Shyf-Fang*
Developing a Two-Dimensional Portfolio Matrix of Innovation and Profitability: Evidence from Solar Energy Firms

INDEX – ASIA CHAPTER

1528 *Chang, Che-Wei, Hung, Yu-Chun, Tsai, Pei-Fen*
Legal Risk Management of Business Owners from the Perspective of Choice-of-Law Rules: A Research Focus On the Impact of Environmental Change on Food Safety

Exploring User Intentions and Behavior Towards Generative AI Adoption in Conversational Commerce

Pao-Jung Chao, Szu-Ying Lai, Chih-Chieh Chang
National Taiwan University of Science and Technology, Taipei, Taiwan

Abstract

Generative Artificial Intelligence (GenAI) is transforming retail from static e-commerce to interactive conversational commerce (C-commerce). Unlike traditional chatbots, which struggle with complex inquiries and produce formulaic responses, GenAI leverages advanced Natural Language Processing (NLP) to enable contextual awareness, personalisation, and empathetic dialogue. This paper argues that GenAI-based chatbots represent a shift toward intelligent knowledge systems capable of fostering organisational resilience and inclusive innovation.

To explore this phenomenon, we integrate three foundational frameworks: the Stimulus-Organism-Response (SOR) framework, the Technology Acceptance Model (TAM), and the Information System Success (ISS) Model. The GenAI chatbot environment serves as the external Stimulus, shaping the user's internal Organism state — comprising TAM-derived perceptions (Ease of Use, Usefulness, Risk, and Enjoyment) and ISS-derived Perceived Trust — which ultimately drives the behavioural Response (intention and actual usage).

The framework was validated through an online survey (N = 283) conducted in Taiwan, featuring a scenario-based interaction involving a hypothetical perfume purchase. Data were analysed using Partial Least Squares Structural Equation Modelling (PLS-SEM) in SmartPLS 4.0. The measurement model demonstrated strong reliability (Cronbach's α: 0.819–0.903; CR > 0.7) and convergent validity (AVE: 0.541–0.816). Seven of nine hypotheses were supported (H1, H2, H4, H5, H7, H8, H9). Behavioural intention significantly predicted actual behaviour ($\beta = 0.746$, $p < 0.001$, $R^2 = 0.556$). Intention was positively influenced by perceived usefulness ($\beta = 0.243$), perceived trust ($\beta = 0.238$), perceived enjoyment ($\beta = 0.211$), and perceived ease of use ($\beta = 0.175$). Perceived trust was driven by information quality ($\beta = 0.390$) and service quality ($\beta = 0.341$). Perceived risk and system quality showed no significant effects, suggesting that technical availability has become a baseline expectation, while trust is anchored in knowledge asset quality — response accuracy, empathy, and personalisation.

These findings support a stratified hybrid implementation strategy: deploying traditional automation for routine administrative tasks while reserving GenAI for high-empathy interactions such as product recommendations. This approach positions chatbots not merely as sales assistants but as proactive knowledge agents that cultivate customer needs and maximise organisational knowledge value.

Keywords: Generative AI, Conversational Commerce, TAM, SOR, ISS Model

Paper type: Academic Research Paper

1 Introduction

Generative Artificial Intelligence (GenAI) has emerged as a disruptive force following the launch of ChatGPT in late 2022, reaching one million users within just five days. Based on Transformer models, GenAI possesses advanced semantic understanding and contextual conversation skills. In the business landscape, this technology is poised to revolutionize conversational commerce (C-commerce), where consumer purchasing intentions are shaped through interactive dialogues.

1.1 Conversational commerce and generative AI landscape

The retail market has evolved from traditional brick-and-mortar stores to digital e-commerce, and subsequently to social commerce driven by messaging apps. C-commerce serves as a direct interface connecting customers and companies through text-based chatbots or voice assistants like Siri and Alexa. While the industry is experiencing rapid growth with a projected multi-billion dollar market value , the driving force behind this expansion is the integration of machine learning and AI to provide personalized engagement.

Unlike traditional models designed for classification, GenAI models—such as the Generative Pre-trained Transformer (GPT) family—can generate fresh, human-like content from extensive datasets. These tools exhibit contextual awareness, self-explanation capabilities, and personalization, distinguishing them from previous language models. Consequently, global retailers are increasingly integrating GenAI to optimize search results and enhance customer service.

1.2 The gap in application and research objectives

Despite the convenience of C-commerce, consumers often remain hesitant to utilize traditional chatbots. Major challenges include a perceived lack of empathy and knowledge, and an inability to handle abstract inquiries such as humor or sarcasm. Conversely, GenAI introduces new challenges, including the potential for "hallucinations" or responses that are out of context, alongside persistent concerns regarding user privacy and trust.

There is currently limited knowledge regarding the specific factors influencing the integration of GenAI-based chatbots into C-commerce. To address this gap, this study integrates the Stimulus-Organism-Response (SOR) framework, the Technology Acceptance Model (TAM), and the DeLone and McLean Information Systems Success (ISS) Model. We aim to develop a novel integrated framework to provide a comprehensive understanding of the causal factors behind behavioral intention and actual usage, while exploring how content reliability, system functionality, and service efficiency affect user trust. By achieving these objectives, this research provides meaningful insights for brand executives and developers to foster stronger customer connections through the strategic implementation of GenAI technology.

2. Literature review

This research formulates a theoretical model integrating the TAM, SOR, and DeLone and McLean ISS frameworks to examine the adoption of Generative-AI-based chatbots.

2.1 Technology acceptance model (TAM)

TAM, originally introduced by Davis (1989), evolved from social psychology theories such as the Theory of Reasoned Action (TRA) and the Theory of Planned Behavior (TPB). While extended models like TAM2 and UTAUT incorporate additional variables such as social influence and effort expectancy, TAM remains a cornerstone for understanding technology adoption. It centers on two primary drivers: perceived usefulness (PU) and perceived ease of use (PE). In the domain of e-commerce, TAM is recognized for its superior flexibility and scalability, providing a reliable framework for investigating user intentions toward AI integration and chatbot services. Consequently, TAM is employed as a foundational model for this study.

2.2 Stimulus–organism–response (SOR) framework

The SOR framework, derived from environmental psychology, posits that consumer behavior is determined by external stimuli that affect internal cognitive and affective states.

- Stimulus: Represents external environmental factors.
- Organism: Refers to the internal psychological processes of the individual.
- Response: The resulting behavioral outcome, such as approach or avoidance.

SOR is widely applied to clarify motivations and experiences in online shopping environments. This study utilizes SOR to investigate how the psychological environment created by GenAI-based chatbots (stimulus) impacts users' internal perceptions (organism) and subsequent intentions (response).

2.3 DeLone and McLean information systems success (ISS) model

The ISS model identifies critical success factors for information systems. While the original 1992 version focused on system and information quality, the updated 2003 model introduced service quality as a third essential dimension.

- System Quality: Evaluates technical success through efficiency and reliability.
- Information Quality: Measures semantic success via accuracy and relevance.
- Service Quality: Assesses overall user fulfillment and professionalism.

Previous research highlights that trust is crucial in bridging these quality dimensions with consumer attitudes. This research intends to fill the gap in how these three quality factors specifically influence users’ trust when utilizing GenAI-based chatbots in C-commerce.

3 Hypotheses Development

This study integrates the SOR, TAM, and ISS models to explore the adoption of GenAI-based chatbots. The SOR framework serves as the foundation: GenAI features act as the Stimulus, influencing users' internal Organism (perceptions derived from TAM and ISS), which ultimately leads to a behavioral Response (intention and actual usage).

3.1 Perceived factors and intentions (TAM & organism)

- H1 (Perceived Ease of Use): Previous research indicates that a user-friendly interface requiring minimal mental effort significantly enhances chatbot adoption intentions.
- H2 (Perceived Usefulness): Studies show that users' belief in a technology's potential to improve performance is a critical predictor of their intent to engage with AI-based services.
- H3 (Perceived Risk): Research suggests that concerns regarding data privacy and information accuracy can negatively impact the pursuit of desired outcomes in C-commerce.
- H4 (Perceived Enjoyment): Beyond performance, the intrinsic pleasure or interest derived from interacting with advanced AI serves as a powerful motivator for technology use.

3.2 Quality dimensions and perceived trust (ISS & organism)

- H5 (Information Quality): High standards of accuracy, relevance, and completeness are essential prerequisites for establishing user trust in conversational agents.
- H6 (System Quality): Technical success, measured through system flexibility, reliability, and response time, is hypothesized to bolster the perceived trustworthiness of the chatbot.
- H7 (Service Quality): The professionalism, promptness, and personalization of the service are critical for fostering positive attitudes and initial trust in a C-commerce context.

3.3 The path to actual behavior (response)

- H8 (Trust): In high-risk online environments, trust serves as a mechanism to reduce uncertainty, thereby positively influencing behavioral intentions.
- H9 (Behavioral Intention & Actual Behavior): Consistent with behavioral theories, a person's rational decision to use a service (intention) is the direct precursor to their actual usage behavior.

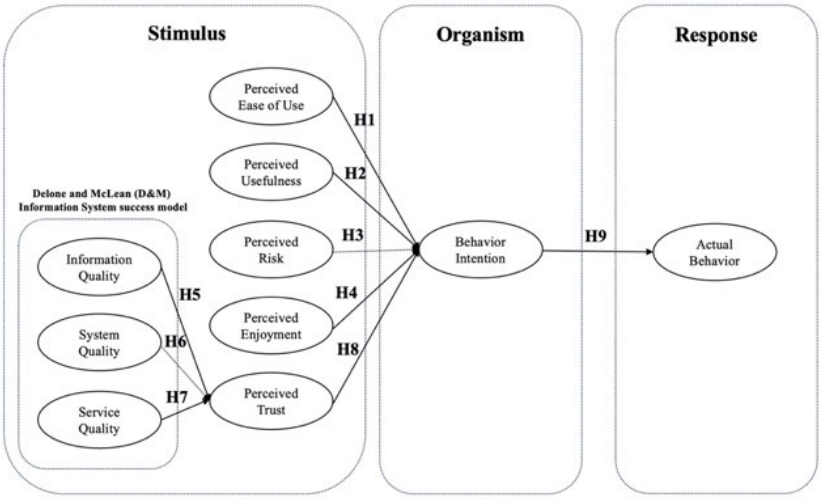


Figure 1 Research Model
Consumers' responses to Generative-AI-based chatbot in conversational commerce.

4 Research Methodology

4.1. Research design and measurement

To empirically validate the interplay between human cognition and artificial intelligence within sustainable digital ecosystems, an online survey was conducted in Taiwan between November and December 2023. The research design followed a three-phase protocol aimed at capturing the nuances of intelligent knowledge in practice. Initially, participants were introduced to the conceptual convergence of conversational commerce and Generative AI (GenAI), ensuring a foundational understanding of how these technologies reshape organizational knowledge flows. This was followed by a scenario-based simulation where participants interacted with a GenAI-driven assistant for a hypothetical perfume purchase, allowing for a direct assessment of AI's role in augmenting decision-making processes. The final phase gathered demographic data alongside assessments of theoretical

variables, using a 7-point Likert scale to capture the depth of participant agreement. To ensure cross-cultural validity, the instrument was provided in Traditional Chinese and English, with linguistic refinements made by native speakers to maintain clarity for both domestic and international respondents.

The measurement framework integrated forty-five items adapted from seminal literature to evaluate the systemic and relational dimensions of human-AI collaboration. Specifically, the study employed Davis's metrics for perceived usefulness, ease of use, and enjoyment to gauge technological acceptance, while the DeLone and McLean model provided the basis for assessing information, system, and service quality. Furthermore, constructs of perceived risk, trust, and behavioral intentions—drawing from Featherman, Pavlou, and Kasilingam—were utilized to analyze the ethical and social implications of AI integration. From a total of 283 valid responses (93 male, 190 female), the analysis demonstrated robust internal consistency, with Cronbach's alpha values ranging from 0.819 to 0.903. These results significantly exceed the reliability threshold of 0.7, confirming that the measurement model offers a reliable platform for exploring how intelligent knowledge systems can foster resilient and sustainable consumer-organizational relationships in an age of complexity.

4.2. Descriptive statistical analysis of samples

Data were collected online in November 2023 via Google Forms. After removing 10 duplicate responses, 283 valid responses were retained (275 Taiwanese; 8 foreign residents). Respondents were predominantly aged 18–25 (64.31%), followed by 26–35 (21.20%) and 36 and above (14.49%). Educationally, 49.82% held a bachelor's degree and 41.70% a master's degree. Most respondents (89.75%) used E-commerce platforms (MOMO, SHOPEE, PCHOME) as their primary shopping channel, and 70.67% shopped online 1–3 times per month. Demographic details are summarised in Table 1.

Regarding descriptive statistics, mean agreement scores across all constructs ranged from 4.37 to 6.11 on a 7-point Likert scale. Perceived Ease of Use (PE: 5.38–5.57), Perceived Usefulness (PU: 5.59–5.84), and System Quality (SSI: 5.31–6.11) received relatively higher scores, while Perceived Risk (PR: 4.42–4.78) and Service Quality (SQ: 4.37–5.25) were comparatively lower. Trust (TT: 4.78–5.16), Perceived Enjoyment (PEN: 5.06–5.14), Information Quality (IQ: 5.31–5.54), Behavioural Intention (BI: 5.15–5.46), and Actual Behaviour (AB: 5.07–5.55) all fell within a moderate-to-high range.

Table 1 Samples statistics: Demographic information of the respondents

Demographic characteristics		Frequency	Percentage
Gender	Male	93	32.86
	Female	190	67.14
Age	18-25	182	64.31
	26-35	60	21.20
	36 and above	41	14.49
Education	High School	19	6.71
	Bachelor	141	49.82
	Master	118	41.70
	PhD	5	1.77
Frequency of Online Shopping per Month	0 times	14	4.95
	1-3 times	200	70.67
	4-6 times	51	18.02
	7 times and above	18	6.36
Platform	Brand Official Website	26	9.19
	E-Commerce Platform	254	89.75
	Social Media	3	1.06
Nationality	TW	275	97.17
	Vietnam	4	1.41
	Indonesia	3	1.06
	Thailand	1	0.35

4.3. Data analysis and results

PLS-SEM was conducted using SmartPLS 4.0, an approach well-suited for complex models with non-normal data (Hair et al., 2017; Willaby et al., 2015; Hair et al., 2009). Following Anderson and Gerbing's two-step procedure, the measurement model was assessed first, followed by structural model analysis.

4.3.1 Measurement model of reliability and convergent validity

As shown in Table 2, all Cronbach's α values (0.819–0.903) and composite reliability (CR) values exceed the 0.7 threshold (Nunnally and Bernstein, 1994), confirming internal consistency. AVE values (0.541–0.816) surpass the 0.5 threshold (Fornell and Larcker, 1981; Anderson and Gerbing, 1988), and all standardised factor loadings fall between 0.610 and 0.970 (t > 1.96), demonstrating convergent validity and indicator reliability. Based on the Fornell–Larcker criterion (Table 3), the minimum square root of AVE (0.857) exceeds the highest inter-construct correlation (0.746), confirming discriminant validity (Fornell and Larcker, 1981). Overall, all constructs meet the requirements for reliability, convergent validity, and discriminant validity, and the analysis proceeds to structural model testing.

Table 2 Reliability and validity results

Constructs	Items	Factor loadings	t-value	Cronbach's α	CR	AVE
Perceived Ease of Use(PE)	PE1 <- PE	0.722	14.816	0.885	0.916	0.688
	PE2 <- PE	0.851	42.143			
	PE3 <- PE	0.881	41.255			
	PE4 <- PE	0.901	60.558			
	PE5 <- PE	0.779	15.286			
Perceived Usefulness(PU)	PU1 <- PU	0.812	32.561	0.902	0.927	0.718
	PU2 <- PU	0.853	41.441			
	PU3 <- PU	0.885	54.738			
	PU4 <- PU	0.873	49.803			
	PU5 <- PU	0.811	28.833			
Perceived Risk(PR)	PR1 <- PR	0.793	3.052	0.842	0.887	0.725
	PR2 <- PR	0.970	3.861			
	PR3 <- PR	0.778	2.909			
Perceived Enjoyment(PEN)	PEN1 <- PEN	0.864	40.946	0.887	0.930	0.816
	PEN2 <- PEN	0.925	93.241			
	PEN3 <- PEN	0.918	75.534			
Information Quality(IQ)	IQC1 <- IQ	0.851	44.199	0.903	0.925	0.673
	IQC2 <- IQ	0.804	28.291			
	IQC3 <- IQ	0.825	30.141			
	IQC4 <- IQ	0.787	21.846			
	IQC5 <- IQ	0.813	30.405			
	IQC6 <- IQ	0.841	36.415			
System Quality(SSI)	SSI1 <- SSI	0.610	11.342	0.868	0.891	0.541
	SSI2 <- SSI	0.704	13.771			
	SSI3 <- SSI	0.730	13.809			
	SSI4 <- SSI	0.773	18.739			
	SSI5 <- SSI	0.801	16.847			
	SSI6 <- SSI	0.772	14.435			
	SSI7 <- SSI	0.744	14.406			
Service Quality(SQ)	SQ1 <- SQ	0.751	21.386	0.890	0.916	0.645
	SQ2 <- SQ	0.840	42.713			
	SQ3 <- SQ	0.812	32.699			
	SQ4 <- SQ	0.825	29.368			
	SQ5 <- SQ	0.796	23.459			
	SQ6 <- SQ	0.791	28.590			
Trust(TT)	TT1 <- TT	0.897	61.246	0.879	0.917	0.735
	TT2 <- TT	0.911	62.717			
	TT3 <- TT	0.819	24.960			
	TT4 <- TT	0.797	29.030			
Behavioral Intention to Use(BI)	BIU1 <- BIU	0.813	30.804	0.826	0.896	0.743
	BIU2 <- BIU	0.880	35.433			
	BIU3 <- BIU	0.891	63.701			
Actual Behavior(AB)	AB1 <- AB	0.875	58.145	0.819	0.891	0.733
	AB2 <- AB	0.887	51.826			
	AB3 <- AB	0.804	23.858			

Table 3 Reliability and validity test results.

	AB	BIU	IQ	PE	PEN	PR	PU	SQ	SSI	TT
AB	0.856									
BIU	0.746	0.862								
IQ	0.517	0.598	0.820							
PE	0.499	0.502	0.473	0.829						
PEN	0.546	0.544	0.513	0.445	0.903					
PR	0.061	0.028	0.091	0.101	0.173	0.852				
PU	0.576	0.580	0.550	0.619	0.542	0.052	0.848			
SQ	0.558	0.651	0.681	0.413	0.524	0.118	0.467	0.803		
SSI	0.556	0.542	0.642	0.578	0.439	0.016	0.572	0.451	0.736	
TT	0.508	0.535	0.637	0.376	0.567	0.129	0.496	0.617	0.427	0.857

Note: The square roots of average variance extracted (AVE) are highlighted in bold.

4.3.2. Structural model

The structural model was evaluated using SmartPLS 4.0 with 5,000 bootstrap subsamples, a bias-corrected accelerated method, and a two-tailed significance level of 0.05. Model fit was confirmed by a GoF of 0.500, exceeding the 0.36 threshold (Cohen, 1988; Tenenhaus et al., 2005).

As shown in Table 4, behavioural intention was significantly influenced by perceived ease of use ($\beta = 0.175$, $p < 0.01$), perceived usefulness ($\beta = 0.243$, $p < 0.01$), perceived enjoyment ($\beta = 0.211$, $p < 0.01$), and perceived trust ($\beta = 0.238$, $p < 0.001$), supporting H1, H2, H4, and H8. Perceived risk ($\beta = -0.07$, $p > 0.05$) was non-significant, thus H3 is not supported. For perceived trust, information quality ($\beta = 0.390$, $p < 0.001$) and service quality ($\beta = 0.341$, $p < 0.001$) were significant predictors (H5, H7 supported), while system quality ($\beta = 0.022$, $p > 0.05$) was not (H6 not supported). Behavioural intention strongly predicted actual behaviour ($\beta = 0.746$, $p < 0.001$), supporting H9.

Table 5 shows moderate predictive ability for Trust and Behavioural Intention ($R^2 = 0.469$) and high predictive ability for Actual Behaviour ($R^2 = 0.556$). Effect sizes indicated a weak influence of Trust on Behavioural Intention ($f^2 = 0.067$) and a strong influence of Behavioural Intention on Actual Behaviour ($f^2 = 1.253$). No multicollinearity issues were detected. Overall, the structural model meets academic standards for explanatory power, predictive ability, and goodness of fit.

Table 4 Path Coefficients.

	Path	β	t-value	p-value	Result
H1:	PE -> BI	0.175	2.659	0.008**	Supported
H2:	PU -> BI	0.243	3.290	0.001**	Supported
H3:	PR -> BI	-0.07	1.120	0.263	Not Supported
H4:	PEN -> BI	0.211	2.974	0.003**	Supported
H5:	IQ -> TT	0.390	4.133	0.000***	Supported
H6:	SSI -> TT	0.022	0.376	0.707	Not Supported
H7:	SQ -> TT	0.341	5.003	0.000***	Supported
H8:	TT -> BI	0.238	3.589	0.000***	Supported
H9:	BI -> AB	0.746	24.538	0.000***	Supported

Note: "*" represents $p < 0.05$, "**" $p < 0.01$, "***" $p < 0.001$

Table 5 Result of R2 , R2 adjusted, F2 , and GoF

		R^2	R^2 adjusted	f^2	GoF
H1:	PE -> BI			0.034	
H2:	PU -> BI			0.056	
H3:	PR -> BI			0.009	
H4:	PEN -> BI			0.048	
H5:	IQ -> TT			0.114	
H6:	SSI -> TT			0.001	
H7:	SQ -> TT	0.469	0.463	0.117	
H8:	TT -> BI	0.469	0.459	0.067	0.500
H9:	BI -> AB	0.556	0.555	1.253	

Note: "*" represents $p < 0.05$, "**" $p < 0.01$, "***" $p < 0.001$

5 Discussion

This study integrates TAM, SOR, and ISS models to examine factors influencing users' behavioural intention and actual behaviour toward GenAI-based chatbots in C-commerce. The model explains 46.9% of variance in behavioural intention and 55.6% in actual behaviour. Seven hypotheses were supported (H1, H2, H4, H5, H7, H8, H9); H3 and H6 were not supported.

Regarding H1, perceived ease of use was significant, consistent with prior research (Mahendar et al., 2023; Jan et al., 2023). GenAI chatbots offer intuitive interfaces that lower interaction barriers, expedite transactions, and strengthen users' intention to engage.

Regarding H2, perceived usefulness was supported, reinforcing TAM's core proposition (Davis, 1989; Mahendar et al., 2023; Pillai and Mukherjee, 2011). Users perceive GenAI chatbots as performance-enhancing tools that streamline the shopping process in C-commerce.

Regarding H3, perceived risk was non-significant, diverging from some prior studies (Salisbury et al., 2001; Ariffin et al., 2018; Wu et al., 2020) but aligning with others (Kasilingam, 2020; Liu et al., 2023). Possible explanations include risk normalisation among experienced online shoppers, positive ease-of-use experiences that offset privacy concerns, and the growing acceptance of AI in daily life. Nonetheless, practitioners should address transaction security and data privacy to broaden adoption.

Regarding H4, perceived enjoyment significantly influenced behavioural intention (Kasilingam, 2020; Teo et al., 1999). Unlike task-oriented traditional chatbots, GenAI systems generate contextually varied and engaging responses, creating an enjoyable conversational environment that motivates continued use.

Regarding H5, information quality positively affected perceived trust (Liu et al., 2023; Teo et al., 2008), as users expect accurate, relevant, and complete responses. GenAI's capacity to deliver personalised, context-aware answers directly strengthens trust perceptions.

Regarding H6, system quality was non-significant, consistent with findings in banking chatbot research (Nguyen et al., 2021). Users appear to treat 24/7 availability and response speed as baseline expectations rather than trust-building attributes, directing trust judgements toward response accuracy, personality, and problem-solving ability.

Regarding H7, service quality significantly enhanced perceived trust (Roh et al., 2022). GenAI's ability to comprehend user needs and respond with empathy elevates perceived service standards beyond what rule-based chatbots can offer.

Regarding H8, perceived trust positively influenced behavioural intention (Harris and Goode, 2004; Pavlou, 2003), confirming its central role in C-commerce adoption. Higher trust reduces users' motivation to switch to human agents, thereby supporting chatbot acceptance.

Regarding H9, behavioural intention was a significant positive predictor of actual behaviour ($\beta = 0.746$). Consistent with SOR theory, when users perceive the chatbot environment as useful, enjoyable, and trustworthy, the resulting internal intention translates directly into actual usage behaviour.

6 Conclusion

This study examined factors influencing consumer behavioural intention and actual behaviour toward GenAI-based chatbots in C-commerce, integrating TAM, ISS, and SOR frameworks. Findings confirm that perceived ease of use, usefulness, enjoyment, and trust drive intention, while information quality and service quality underpin trust. System quality and perceived risk were non-significant, suggesting users treat technical availability as a baseline expectation and have normalised AI-related risks through prior online shopping experience.

Practically, we advocate a stratified hybrid strategy: deploying traditional rule-based automation for routine tasks (order updates, address changes) while reserving GenAI for high-empathy interactions such as product recommendations. This positions GenAI chatbots not merely as sales assistants but as proactive agents that cultivate customer needs, enhance brand trust, and bridge the gap between automated service and human customer support.

This study has several limitations. The sample skews toward young, female participants in Taiwan, and the scope is limited to text-based GenAI. Future research should broaden demographic diversity, explore image- and voice-based AI applications, examine privacy-related concerns, and consider alternative theoretical frameworks to further advance understanding of AI adoption in digital commerce.

7 Ethics declaration

This study did not require ethical clearance as it involved anonymous online survey data with no sensitive personal information collected.

References

Anderson, J.C. and Gerbing, D.W. (1988) Structural equation modeling in practice: a review and recommended two-step approach, Psychological Bulletin, Vol. 103, No. 3, pp. 411–423.

Ariffin, S.K., Mohan, T.R.M. and Goh, Y. (2018) Influence of consumers' perceived risk on consumers' online purchase intention, Journal of Research in Interactive Marketing, Vol. 12, No. 3, pp. 309–327.

Cohen, J. (1988) Statistical Power Analysis for the Behavioral Sciences, 2nd ed., Lawrence Erlbaum Associates, Hillsdale, NJ.

Davis, F. (1989) Perceived usefulness, perceived ease of use, and user acceptance of information technology, MIS Quarterly, Vol. 13, No. 3, pp. 319–340.

Davis, F.D., Bagozzi, R.P. and Warshaw, P.R. (1992) Extrinsic and intrinsic motivation to use computers in the workplace, Journal of Applied Social Psychology, Vol. 22, No. 14, pp. 1111–1132.

DeLone, W.H. and McLean, E.R. (1992) Information systems success: the quest for the dependent variable, Information Systems Research, Vol. 3, No. 1, pp. 60–95.

DeLone, W.H. and McLean, E.R. (2003) The DeLone and McLean model of information systems success: a ten-year update, Journal of Management Information Systems, Vol. 19, No. 4, pp. 9–30.

Featherman, M. and Pavlou, P.A. (2003) Predicting e-services adoption: a perceived risk facets perspective, International Journal of Human-Computer Studies, Vol. 59, No. 4, pp. 451–474.

Fornell, C. and Larcker, D.F. (1981) Evaluating structural equation models with unobservable variables and measurement error, Journal of Marketing Research, Vol. 18, No. 1, pp. 39–50.

Hair, J.F., Hult, G.T.M., Ringle, C.M. and Sarstedt, M. (2017) A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM), 2nd ed., Sage Publications, Thousand Oaks, CA.

Hair, J.F., Black, W.C., Babin, B.J. and Anderson, R.E. (2009) Multivariate Data Analysis, 7th ed., Prentice Hall, Upper Saddle River.

Harris, L.C. and Goode, M.M.H. (2004) The four levels of loyalty and the pivotal role of trust: a study of online service dynamics, Journal of Retailing, Vol. 80, No. 2, pp. 139–158.

Jan, I.U., Ji, S. and Kim, C. (2023) What (de)motivates customers to use AI-powered conversational agents for shopping? The extended behavioral reasoning perspective, Journal of Retailing and Consumer Services, Vol. 75, 103440.

Kasilingam, D. (2020) Understanding the attitude and intention to use smartphone chatbots for shopping, Technology in Society, Vol. 62, 101280.

Liu, Y., Hu, B., Yan, W. and Zhi, L. (2023) Can chatbots satisfy me? A mixed-method comparative study of satisfaction with task-oriented chatbots in mainland China and Hong Kong, Computers in Human Behavior, Vol. 143, 107716.

Nguyen, D.T., Chiu, Y.H. and Le, H.D. (2021) Determinants of continuance intention towards banks' chatbot services in Vietnam: a necessity for sustainable development, Sustainability, Vol. 13, No. 14, 7625.

Nunnally, J.C. and Bernstein, I.H. (1994) Psychometric Theory, 3rd ed., McGraw-Hill, New York.

Lowry, P.B. and Gaskin, J. (2014) Partial least squares (PLS) structural equation modeling (SEM) for building and testing behavioral causal theory: when to choose it and how to use it, IEEE Transactions on Professional Communication, Vol. 57, No. 2, pp. 123–146.

Pavlou, P.A. (2003) Consumer acceptance of electronic commerce: integrating trust and risk with the technology acceptance model, International Journal of Electronic Commerce, Vol. 7, No. 3, pp. 101–134.

Pillai, A. and Mukherjee, J. (2011) User acceptance of hedonic versus utilitarian social networking web sites, Journal of Indian Business Research, Vol. 3, No. 3, pp. 180–191.

Roh, T., Yang, Y.S., Xiao, S. and Park, B.I. (2022) What makes consumers trust and adopt fintech? An empirical investigation in China, Electronic Commerce Research, Vol. 24, No. 1, pp. 3–35.

Salisbury, W.D., Pearson, R.A., Pearson, A.W. and Miller, D.W. (2001) Perceived security and World Wide Web purchase intention, Industrial Management and Data Systems, Vol. 101, No. 4, pp. 165–177.

Tenenhaus, M., Vinzi, V.E., Chatelin, Y.M. and Lauro, C. (2005) PLS path modeling, Computational Statistics and Data Analysis, Vol. 48, pp. 159–205.

Teo, T.S.H., Lim, V.K.G. and Lai, R.Y. (1999) Intrinsic and extrinsic motivation in Internet usage, Omega, Vol. 27, No. 1, pp. 25–37.

Teo, T.S.H., Srivastava, S.C. and Jiang, L. (2008) Trust and electronic government success: an empirical study, Journal of Management Information Systems, Vol. 25, No. 3, pp. 99–132.

Willaby, H.W., Costa, D., Burns, B.D., MacCann, C. and Roberts, R.D. (2015) Testing complex models with small sample sizes: a historical overview and empirical demonstration of what partial least squares (PLS) can offer differential psychology, Personality and Individual Differences, Vol. 84, pp. 73–78.

Wu, I., Chiu, M. and Chen, K. (2020) Defining the determinants of online impulse buying through a shopping process of integrating perceived risk, expectation-confirmation model, and flow theory issues, International Journal of Information Management, Vol. 52, 102099.

Exploring Federated Learning Adoption in Competitive and Collaborative Interorganizational Environments: A TOE-S Perspective

Tzu-Hsuan Wang, Yi-Shan Chen, Chih-Chieh Chang
Management Undergraduate Program, National Taiwan University of Science and Technology, Taipei, Taiwan

Abstract

In the current era of rapid digital transformation, financial institutions navigate a profound strategic paradox: the urgent requirement for collaborative Artificial Intelligence intelligence to combat increasingly sophisticated financial crimes versus increasingly stringent data privacy mandates that result in fragmented data silos. Federated Learning (FL) has emerged as a transformative innovation to resolve this privacy-utility trade-off, enabling decentralized machine learning where model training occurs locally on edge devices. This mechanism allows organizations to bridge critical information gaps without the risky transfer of sensitive raw data. However, the diffusion of FL within coopetitive environments—where rivals must simultaneously compete for market share and cooperate against systemic threats—remains an under-researched domain, particularly regarding the synergy between automated systems and human cognitive judgment.

This study addresses this theoretical and practical gap by investigating the socio-technical factors influencing Federated Learning adoption within the Taiwanese banking sector for collaborative fraud detection alliances. We propose an extended Technology-Organization-Environment-Support (TOE-S) framework, integrating internal leadership vision and external regulatory clarity as pivotal catalysts for innovation. A quantitative research design was employed, collecting 188 valid responses from banking professionals in Taiwan involved in anti-fraud and risk management operations. Data were analyzed using Partial Least Squares Structural Equation Modeling (PLS-SEM) to evaluate the relationships between the constructs and validate the theoretical model.

The empirical results demonstrate that the proposed theoretical model possesses strong explanatory power for both adoption intention and actual usage behavior. The analysis reveals that Top Management Support and Regulatory Support are the primary predictors that exert a significant positive impact on the intention to adopt this technology. Interestingly, traditional technological hurdles, such as compatibility and complexity, as well as external competitive pressures, were found to be statistically non-significant in this context. This suggests that in highly regulated ecosystems, institutional trust and leadership mandate supersede technical considerations. The study offers a strategic roadmap for fostering resilient and sustainable financial organizations through responsible governance and inclusive innovation. The findings underscore that Federated Learning adoption should be treated as a core strategic pillar for organizational resilience rather than a mere technical Information Technology (IT) project.

Keywords: Federated Learning, Intelligent Knowledge, TOE-S Framework, Financial Fraud Detection, Coopetition

Paper type: Academic Research Paper

1 Introduction

1.1 Research Background and Intelligent Knowledge

The rapid advancement of Artificial Intelligence (AI) has empowered enterprises to automate repetitive tasks and enhance operational efficiency. In the financial sector, the synergy between human cognition and AI tools has become a cornerstone for sustainable growth. However, this evolution faces a critical bottleneck known as the Data Silo problem. Traditional machine learning requires consolidating training data into a single center, a process fraught with data leakage risks and governed by stringent regulations such as the General Data Protection Regulation (GDPR) and Taiwan's Personal Data Protection Act.

Federated Learning (FL) offers a paradigm-shifting solution by allowing model training to occur at the data source and sharing only model parameters, preserving privacy while building collective intelligence. In the context of modern knowledge management, FL represents a shift from knowledge ownership to knowledge orchestration, allowing banks to manage systemic risks through inclusive AI innovation. In Taiwan, financial fraud

cases have shown a significant upward trend, growing by nearly 30% in 2023 alone, according to the National Police Agency. This underscores the urgent need for collaborative resilience. The establishment of the Eagle Eye Anti-Fraud Alliance in 2024 represents Taiwan's first large-scale attempt at using FL for inter-bank collaboration. However, many institutions remain hesitant to adopt FL due to the complex nature of "coopetition"—a strategic state where banks must cooperate for systemic safety while remaining market rivals. This study applies an extended Technology-Organization-Environment-Support (TOE-S) framework to understand the drivers of FL adoption in this unique landscape.

2 Literature Review

2.1 Overview and Taxonomy of Federated Learning

Federated Learning (FL) is a decentralized machine learning paradigm that operates on the fundamental principle of "bringing the model to the data". The standard process involves a central coordinator distributing a global model to participants, who then perform local training on their private datasets. Only model parameters (gradients) are returned to the server for aggregation, and this cycle repeats until the model achieves optimal convergence. This mechanism effectively addresses the strategic paradox of enhancing Artificial Intelligence (AI) intelligence while adhering to stringent data privacy mandates like the General Data Protection Regulation (GDPR).

In interorganizational contexts, FL is primarily categorized into two scenarios based on data distribution. Horizontal Federated Learning (HFL) is utilized when participants share overlapping features but possess different samples, such as multiple banks collaborating to detect fraud using similar transaction record formats for different customer bases. Conversely, Vertical Federated Learning (VFL) applies to situations where participants have overlapping samples but different features, such as a bank collaborating with an e-commerce platform to assess a shared customer's creditworthiness through complementary datasets.

2.2 Theoretical Framework: From TOE to TOE-S

The Technology-Organization-Environment (TOE) framework (Tornatzky and Fleischer, 1990) is a well-established baseline for analyzing Information Technology (IT) adoption, evaluating the process through technological, organizational, and environmental contexts. While the original TOE model is robust, it often overlooks the critical role of institutional guidance in highly sensitive and regulated industries like banking. To better capture the realities of the Taiwanese financial ecosystem, this study proposes an extended TOE-S framework. By integrating the Support (S) dimension—encompassing User Support, Top Management Support, and Regulatory Support—the model provides a more comprehensive analysis of how internal leadership and external legal clarity act as the true catalysts for innovation in "coopetitive" environments.

3 Research Model and Hypothesis Development

The proposed model evaluates ten constructs across four dimensions to predict Intention to Adopt (IA) and Actual Use (AU) (See Figure 1).

Figure 1: Research Model (TOE-S Framework)

Technological and Organizational Contexts: Compatibility (COMP) reflects the alignment with an organization's values and infrastructure, while Complexity (COMX) involves the perceived implementation difficulty, particularly in complex data processes. We also consider Financial Cost (FC) for AI infrastructure and Firm Size (FS) as a proxy for resource abundance. Larger banks with higher technical compatibility are more likely to adopt FL, while high complexity and high initial costs act as significant barriers. Thus, we propose: H1: COMP(+), H2: COMX(-), H3: FC(-), and H4: FS(+) impact IA.

Environmental and Support Contexts: Competitive Pressure (CP) and Environmental Uncertainty (EU) drive innovation for survival, while Security and Privacy Regulatory Concerns (SPC) may hinder it due to compliance anxiety. Furthermore, internal User Support (US) and Top Management Support (TMS) are critical for strategic alignment, while external Regulatory Support (RS) provides institutional legitimacy. We suggest: H5: CP(+), H6: EU(+), H7: SPC(-), H8: US(+), H9: TMS(+), and H10: RS(+) impact IA.

Behavioral Outcome: Consistent with technology acceptance literature, adoption intention is a robust and direct precursor to implementation. We hypothesize H11: IA has a positive impact on AU.

4 Research Methodology

4.1 Research Design and Context

To empirically validate the TOE-S framework, this study adopted a cross-sectional quantitative design. Measurement items were adapted from validated scales (Tiwari et al., 2023; Wael Al-khatib, 2023) and refined through a pre-test with 17 experts. All indicators were measured using a 7-point Likert scale, which provides higher data variance and precise respondent mapping (Vagias, 2006). Data were collected from August to October 2024 via an anonymous online survey. To fulfill Corporate Social Responsibility (CSR) goals, a donation was made to the TFCF for each valid response, yielding 188 valid samples (89% response rate).

The demographic profile of the respondents reflects a diverse cross-section of the industry, as presented in Table 1. Notably, 48.40% are under 30 years old and 36.17% possess an Information Technology (IT) background. For data analysis, we utilized Partial Least Squares Structural Equation Modeling (PLS-SEM) via SmartPLS 4.0, following criteria by Hair et al. (2021), using a bootstrapping procedure with 5,000 resamples to ensure statistical power.

Table 1: Demographics of the respondents (n=188)

Demographics variable	**Categories**	**Frequency**	**Percent**
Gender	Male	88	46.81%
	Female	100	53.19%
Age	Less than 30 years old	91	48.40%
	31-40 years old	36	19.15%
	41-50 years old	33	17.55%
	Above 51 years old	28	14.89%
Education	Secondary school or below	5	2.66%
	Undergraduate degree	103	54.79%
	Graduate degree	73	38.83%
	Doctorate degree	7	3.72%
IT background	Posses	68	36.17%
	Lack of	120	63.83%
JOB title	Staff	129	68.62%
	Senior staff	20	10.64%
	Basic supervisor	19	10.11%
	Mid-level supervisor	4	2.13%
	Executive-level supervisor	16	8.51%
Work Experience	Less than or equal to 5 years	94	50.00%
	6-10 years	25	13.30%
	11-15 years	17	9.04%
	16-20 years	22	11.70%
	More than 21 years	30	15.96%
Demographics variable	Categories	Frequency	Percent
Firm size	500 people or less	85	45.21%
	Between 501 and 1000 people	28	14.89%
	Between 1001 and 3000 people	23	12.23%
	Between 3001 and 5000 people	9	4.79%
	Between 5001 and 7000 people	9	4.79%
	Between 7001 and 9000 people	9	4.79%
	More than 9001 people	25	13.30%
Number of IT employees	50 people or less	98	52.13%
	Between 51 and 100 people	29	15.43%
	Between 101 and 200 people	23	12.23%
	Between 201 and 400 people	17	9.04%
	More than 401 people	21	11.17%

5 Results

Reliability and validity were evaluated using criteria from Hair et al. (2011). The model demonstrated high reliability (CA and CR > 0.7) and convergent validity (AVE > 0.5). Discriminant validity was confirmed via the HTMT criterion (Henseler et al., 2015), with all values below 0.85 (See Table 2 and Table 3).

Table 2: Convergent Validity Results

Constructs	Indicators	Outer Loadings	Cronbach's alpha	Composite reliability (CR)	Average variance extracted (AVE)
AU	AU1	0.923	0.890	0.890	0.820
	AU2	0.918			
	AU3	0.875			
COMP	COMP1	0.734	0.771	0.790	0.592
	COMP2	0.796			
	COMP3	0.837			
	COMP4	0.704			
COMX	COMX1	0.810	0.797	0.824	0.621
	COMX2	0.865			
	COMX3	0.772			
	COMX4	0.695			
CP	CP1	0.782	0.776	0.881	0.681
	CP2	0.790			
	CP3	0.898			
EU	EU1	0.841	0.736	0.768	0.646
	EU2	0.790			
	EU3	0.779			
FC	FC1	0.961	0.916	0.916	0.922
	FC2	0.96			
FS	FS1	0.926	0.915	0.923	0.855
	FS2	0.938			
	FS3	0.908			
IA	IA1	0.920	0.893	0.893	0.824
	IA2	0.921			
	IA3	0.881			
RS	RS1	0.918	0.770	0.855	0.687
	RS2	0.908			
	RS3	0.628			
SPC	SPC1	0.928	0.754	0.810	0.799
	SPC2	0.858			
TMS	TMS1	0.891	0.885	0.887	0.813
	TMS2	0.941			
	TMS3	0.871			
US	US1	0.912	0.800	0.885	0.710
	US2	0.686			
	US3	0.910			

Table 3: Discriminant Validity (HTMT)

	AU	COMP	COMX	CP	EU	FC	FS	IA	RS	SPC	TMS	US
AU												
COMP	0.746											
COMX	0.503	0.522										
CP	0.659	0.596	0.182									
EU	0.249	0.282	0.186	0.192								
FC	0.034	0.157	0.392	0.356	0.093							
FS	0.530	0.520	0.458	0.277	0.125	0.343						
IA	0.812	0.676	0.420	0.453	0.281	0.160	0.436					
RS	0.491	0.559	0.225	0.368	0.392	0.217	0.264	0.739				
SPC	0.232	0.232	0.605	0.230	0.254	0.353	0.166	0.242	0.124			
TMS	0.812	0.775	0.526	0.560	0.225	0.097	0.444	0.840	0.640	0.310		
US	0.693	0.795	0.698	0.351	0.272	0.235	0.428	0.651	0.562	0.436	0.816	

The model explained 64.4% of the variance in adoption intention (R^2= 0.644) and 52.8% in actual use (R^2= 0.528). Bootstrapping results confirmed that Top Management Support (β = 0.474, $p < 0.001$, supporting H9) and Regulatory Support (β = 0.299, $p < 0.001$, supporting H10) are the primary predictors. Furthermore, IA strongly influenced AU (β = 0.726, $p < 0.001$, supporting H11). Hypotheses H1-H8 were not supported ($p > 0.05$\$), suggesting strategic factors outweigh technical drivers (See Figure 2).

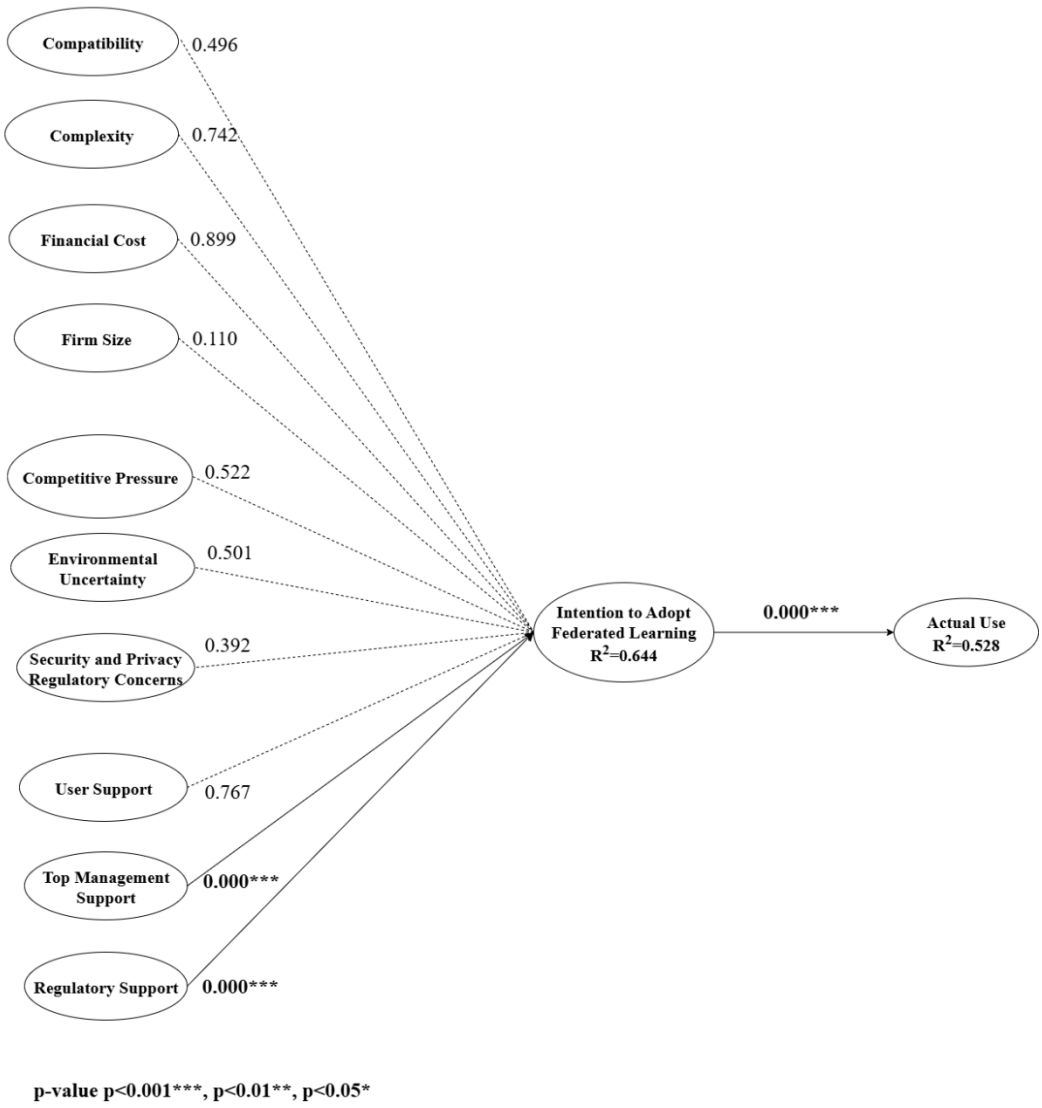


Figure 2: Structural Model Results

6 Discussion

The insignificance of Compatibility (H1) and Complexity (H2) suggests Technical-Strategic Decoupling: banking professionals view technical hurdles as solvable details once a strategic mandate is issued. Moreover, the non-significance of Competitive Pressure (H5) indicates that in Taiwan, fraud detection is viewed as a systemic responsibility rather than a source of competitive advantage. The dominance of Top Management Support and Regulatory Support confirms that innovation in coopetitive landscapes is driven by internal political will and external legal safety.

7 Implications and Conclusion

FL should be framed as a core pillar of Environmental, Social, and Governance (ESG) strategy and organizational resilience. Our findings advocate for Enabled Governance—moving beyond restrictive oversight toward fostering hybrid knowledge ecosystems through regulatory sandboxes. Federated Learning represents a critical solution for the privacy-collaboration paradox, building the collective resilience necessary to thrive in a complex digital landscape.

8 Acknowledgements

The authors express their deepest gratitude to Professor Chih-Chieh Chang for his profound guidance and trust. We also thank the members of the Fintech and AI Lab at NTUST for their support.

9 Ethics Declaration

This study did not involve sensitive human subjects or medical data; therefore, formal ethical review or Institutional Review Board (IRB) approval was not required.

10 AI Declaration

The authors utilized a Large Language Model (LLM) for the structural expansion and linguistic refinement of this manuscript. All statistical data, path coefficients, and theoretical arguments were manually verified and validated by the researchers to ensure factual accuracy and academic rigor.

References

Ahmadi, H., Nilashi, M., Shahmoradi, L., & Ibrahim, O. (2017). Hospital information system adoption: expert perspectives on an adoption framework for Malaysian public hospitals. Computers in Human Behavior, 67, 161–189.

Ahmadi, H., Nilashi, M., Shahmoradi, L., Ibrahim, O., Sadoughi, F., Alizadeh, M., & Alizadeh, A. (2018). The moderating effect of hospital size on inter and intra-organizational factors of Hospital Information System adoption. Technological Forecasting and Social Change, 134, 124–149.

Aibinu, A. A., & Al-Lawati, A. M. (2010). Using PLS-SEM technique to model construction organizations' willingness to participate in e-bidding. Automation in Construction, 19(6), 714–724.

Awosika, T., Shukla, R. M., & Pranggono, B. (2024). Transparency and privacy: the role of explainable AI and Federated Learning in financial fraud detection. IEEE Access, 12, 64551–64560. IEEE Access.

Azman Ong, M. H., Yusri, M. Y., & Ibrahim, N. S. (2023). Use and behavioural intention using digital payment systems among rural residents: Extending the UTAUT-2 model. Technology in Society, 74, 102305.

Bi, X., Gupta, A., & Yang, M. (2024). Understanding partnership formation and repeated contributions in Federated Learning: an analytical investigation. Management Science, 70(8), 4974–4994.

Chaudhuri, R., Singh, B., Agrawal, A. K., Chatterjee, S., Gupta, S., & Mangla, S. K. (2024). A TOE-DCV approach to green supply chain adoption for sustainable operations in the semiconductor industry. International Journal of Production Economics, 275, 109327.

Chen, Y., Yin, Y., Browne, G. J., & Li, D. (2019). Adoption of building information modeling in Chinese construction industry: The technology-organization-environment framework. Engineering, Construction and Architectural Management, 26(9), 1878–1898.

Chittipaka, V., Kumar, S., Sivarajah, U., Bowden, J. L.-H., & Baral, M. M. (2023). Blockchain technology for supply chains operating in emerging markets: An empirical examination of technology-organization-environment (TOE) framework. Annals of Operations Research, 327(1), 465–492.

Cruz-Jesus, F., Pinheiro, A., & Oliveira, T. (2019). Understanding CRM adoption stages: Empirical analysis building on the TOE framework. Computers in Industry, 109, 1–13.

Dayan, I., Roth, H. R., Zhong, A., Harouni, A., Gentili, A., Abidin, A. Z., Liu, A., Costa, A. B., Wood, B. J., Tsai, C.-S., Wang, C.-H., Hsu, C.-N., Lee, C. K., Ruan, P., Xu, D., Wu, D., Huang, E., Kitamura, F. C., Lacey, G., … Li, Q. (2021). Federated learning for predicting clinical outcomes in patients with COVID-19. Nature Medicine, 27(10), 1735–1743.

Duong, C. D., Vu, T. N., & Ngo, T. V. N. (2023). Applying a modified technology acceptance model to explain higher education students' usage of ChatGPT: A serial multiple mediation model with knowledge sharing as a moderator. The International Journal of Management Education, 21(3), 100883.

Durrant, A., Markovic, M., Matthews, D., May, D., Enright, J., & Leontidis, G. (2022). The role of cross-silo federated learning in facilitating data sharing in the agri-food sector. Computers and Electronics in Agriculture, 193, 106648.

Elghdban, M. G., Azmy, N. B., Zulkiple, A. B., & Al-Sharafi, M. A. (2020). Factors affecting the adoption of advanced IT with specific emphasis on building information modeling based on TOE framework: A systematic review. International Journal of Advanced Science and Technology, 29(4), 3314-3333.

Esmaeilzadeh, P. (2024). Challenges and strategies for wide-scale artificial intelligence (AI) deployment in healthcare practices: A perspective for healthcare organizations. Artificial Intelligence in Medicine, 151, 102861.

Gangwar, H., Date, H., & Ramaswamy, R. (2015). Understanding determinants of cloud computing adoption using an integrated TAM-TOE model. Journal of Enterprise Information Management, 28(1), 107–130.

Guan, W., Ding, W., Zhang, B., Verny, J., & Hao, R. (2023). Do supply chain related factors enhance the prediction accuracy of blockchain adoption? A machine learning approach. Technological Forecasting and Social Change, 192, 122552.

Gupta, S., Ghardallou, W., Pandey, D. K., & Sahu, G. P. (2022). Artificial intelligence adoption in the insurance industry: Evidence using the technology–organization–environment framework. Research in International Business and Finance, 63, 101757.

Gutierrez, A., Boukrami, E., & Lumsden, R. (2015). Technological, organisational and environmental factors influencing managers' decision to adopt cloud computing in the UK. Journal of Enterprise Information Management, 28(6), 788–807.

Hair, J. F., Ringle, C. M., & Sarstedt, M. (2011). PLS-SEM: indeed a silver bullet. Journal of Marketing Theory and Practice, 19(2), 139–152.

Hair Jr., J. F., Hult, G. T. M., Ringle, C. M., Sarstedt, M., Danks, N. P., & Ray, S. (2021). Partial Least Squares Structural Equation Modeling (PLS-SEM) Using R: A Workbook. Springer Nature.

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. Journal of the Academy of Marketing Science, 43(1), 115–135.

Hsu, C.-L., & Lin, J. C.-C. (2016). Factors affecting the adoption of cloud services in enterprises. Information Systems and E-Business Management, 14(4), 791–822.

Ilin, V., Ivetić, J., & Simić, D. (2017). Understanding the determinants of e-business adoption in ERP-enabled firms and non-ERP-enabled firms: A case study of the Western Balkan Peninsula. Technological Forecasting and Social Change, 125, 206–223.

Iranmanesh, M., Lim, K. H., Foroughi, B., Hong, M. C., & Ghobakhloo, M. (2022). Determinants of intention to adopt big data and outsourcing among SMEs: Organisational and technological factors as moderators. Management Decision, 61(1), 201–222.

Islam, M., Tamanna, A. K., & Islam, S. (2024). The path to cashless transaction: A study of user intention and attitudes towards quick response mobile payments. Heliyon, 10(15), e35302.

Ivanov, S., Soliman, M., Tuomi, A., Alkathiri, N. A., & Al-Alawi, A. N. (2024). Drivers of generative AI adoption in higher education through the lens of the Theory of Planned Behaviour. Technology in Society, 77, 102521.

Jeyaraj, A., Dwivedi, Y. K., & Venkatesh, V. (2023). Intention in information systems adoption and use: Current state and research directions. International Journal of Information Management, 73, 102680.

Kalloori, S., & Srivastava, A. (2024). Towards cross-silo federated learning for corporate organizations. Knowledge-Based Systems, 289, 111501.

Katebi, A., Homami, P., & Najmeddin, M. (2022). Acceptance model of precast concrete components in building construction based on Technology Acceptance Model (TAM) and Technology, Organization, and Environment (TOE) framework. Journal of Building Engineering, 45, 103518.

Kumar, A., & Shankar, A. (2024). Building a sustainable future with enterprise metaverse in a data-driven era: A technology-organization-environment (TOE) perspective. Journal of Retailing and Consumer Services, 81, 103986.

Kumbhare, S., B.Kathole, A., & Shinde, S. (2023). Federated learning aided breast cancer detection with intelligent Heuristic-based deep learning framework. Biomedical Signal Processing and Control, 86, 105080.

Lai, H.-M., Lin, I.-C., & Tseng, L.-T. (2014). High-Level managers' considerations for RFID adoption in hospitals: An empirical study in Taiwan. Journal of Medical Systems, 38(2), 3.

Lam, T., Cho, V., & Qu, H. (2007). A study of hotel employee behavioral intentions towards adoption of information technology. International Journal of Hospitality Management, 26(1), 49–65.

Li, H., Gupta, A., Zhang, J., & Sarathy, R. (2014). Examining the decision to use standalone personal health record systems as a trust-enabled fair social contract. Decision Support Systems, 57, 376–386.

Li, W., Yigitcanlar, T., Erol, I., & Liu, A. (2021). Motivations, barriers and risks of smart home adoption: From systematic literature review to conceptual framework. Energy Research & Social Science, 80, 102211.

Lin, C., Lin, I.-C., & Roan, J. (2012). Barriers to physicians' adoption of healthcare information technology: An empirical study on multiple hospitals. Journal of Medical Systems, 36(3), 1965–1977.

Lin, H.-F. (2014). Understanding the determinants of electronic supply chain management system adoption: Using the technology–organization–environment framework. Technological Forecasting and Social Change, 86, 80–92.

Liu, Y., Yu, J. J. Q., Kang, J., Niyato, D., & Zhang, S. (2020). Privacy-Preserving traffic flow prediction: A Federated Learning approach. IEEE Internet of Things Journal, 7(8), 7751–7763. IEEE Internet of Things Journal.

Low, C., Chen, Y., & Wu, M. (2011). Understanding the determinants of cloud computing adoption. Industrial Management & Data Systems, 111(7), 1006–1023.

Maroufkhani, P., Tseng, M.-L., Iranmanesh, M., Ismail, W. K. W., & Khalid, H. (2020). Big data analytics adoption: Determinants and performances among small to medium-sized enterprises. International Journal of Information Management, 54, 102190.

McMahan, B., Moore, E., Ramage, D., Hampson, S., & Arcas, B. A. y. (2017). Communication-Efficient learning of deep networks from decentralized data. Proceedings of the 20th International Conference on Artificial Intelligence and Statistics, 1273–1282.

Ministry of the Interior, National Police Agency. (2022). 111年第18週警政統計通報：110年詐欺案件概況 [Week 18 Police Statistics Bulletin 2022: Overview of Fraud Cases in 2021]. National Police Agency.

Ministry of the Interior, National Police Agency. (2024). 113年第14週警政統計通報：112年詐欺案件概況 [Week 14 Police Statistics Bulletin 2024: Overview of Fraud Cases in 2023]. National Police Agency.

Mothukuri, V., Parizi, R. M., Pouriyeh, S., Huang, Y., Dehghantanha, A., & Srivastava, G. (2021). A survey on security and privacy of federated learning. Future Generation Computer Systems, 115, 619–640.

Motie, S., & Raahemi, B. (2024). Financial fraud detection using graph neural networks: A systematic review. Expert Systems with Applications, 240, 122156.

Nam, K., Dutt, C. S., Chathoth, P., Daghfous, A., & Khan, M. S. (2021). The adoption of artificial intelligence and robotics in the hotel industry: Prospects and challenges. Electronic Markets, 31(3), 553–574.

Oliveira, T., Thomas, M., & Espadanal, M. (2014). Assessing the determinants of cloud computing adoption: An analysis of the manufacturing and services sectors. Information & Management, 51(5), 497–510.

Pan, M. J., & Jang, W. Y. (2008). Determinants of the adoption of enterprise resource planning within the technology-organization-environment framework: Taiwan's communications industry. Journal of Computer information systems, 48(3), 94-102.

Peng, W., Gao, J., Chen, Y., & Wang, S. (2024). Bridging data barriers among participants: Assessing the potential of geoenergy through federated learning. Applied Energy, 367, 123293.

Rauniyar, A., Hagos, D. H., Jha, D., Håkegård, J. E., Bagci, U., Rawat, D. B., & Vlassov, V. (2024). Federated Learning for medical applications: A taxonomy, current trends, challenges, and future research directions. IEEE Internet of Things Journal, 11(5), 7374–7398. IEEE Internet of Things Journal.

Rogers, E. M., Singhal, A., & Quinlan, M. M. (2014). Diffusion of innovations. In An integrated approach to communication theory and research (pp. 432-448). Routledge.

Rouhani, S., Ashrafi, A., Ravasan, A. Z., & Afshari, S. (2018). Business intelligence systems adoption model: An empirical investigation. Journal of Organizational and End User Computing (JOEUC), 30(2), 43–70.

Salleh, K. A., & Janczewski, L. (2016). Adoption of Big Data Solutions: A study on its security determinants using Sec-TOE Framework.

Setiyani, L., & Rostiani, Y. (2021). Analysis of E-Commerce adoption by SMEs using the Technology - Organization - Environment (TOE) model: A case study in Karawang, Indonesia. International Journal of Science, Technology & Management, 2, 1113–1132.

Shiranthika, C., Saeedi, P., & Bajić, I. V. (2023). Decentralized learning in healthcare: A review of emerging techniques. IEEE Access, 11, 54188–54209. IEEE Access.

Sun, S., Cegielski, C. G., Jia, L., & Hall, D. J. (2018). Understanding the factors affecting the organizational adoption of big data. Journal of computer information systems, 58(3), 193-203.

Tang, Z., Wong, H.-S., & Yu, Z. (2024). Privacy-Preserving Federated Learning with domain adaptation for multi-disease ocular disease recognition. IEEE Journal of Biomedical and Health Informatics, 28(6), 3219–3227. IEEE Journal of Biomedical and Health Informatics.

Tenenhaus, M., Vinzi, V. E., Chatelin, Y.-M., & Lauro, C. (2005). PLS path modeling. Computational Statistics & Data Analysis, 48(1), 159–205.

Tiwari, A. K., Marak, Z. R., Paul, J., & Deshpande, A. P. (2023). Determinants of electronic invoicing technology adoption: Toward managing business information system transformation. Journal of Innovation & Knowledge, 8(3), 100366.

Tornatzky, L. G., Fleischer, M., & Chakrabarti, A. K. (1990). The Processes of Technological Innovation. Lexington Books.

Towner, L. (2024, December 11). Google cloud and swift pioneer advanced AI and Federated Learning tech to help combat payments fraud. FF News | Fintech Finance.

Venkata Krishna Reddy, V., Vijaya Kumar Reddy, R., Siva Krishna Munaga, M., Karnam, B., Maddila, S. K., & Sekhar Kolli, C. (2024). Deep learning-based credit card fraud detection in federated learning. Expert Systems with Applications, 255, 124493.

Wael AL-khatib, A. (2023). Drivers of generative artificial intelligence to fostering exploitative and exploratory innovation: A TOE framework. Technology in Society, 75, 102403.

Wang, X., Garg, S., Lin, H., Hu, J., Kaddoum, G., Jalil Piran, Md., & Hossain, M. S. (2022). Toward accurate anomaly detection in industrial internet of things using hierarchical Federated Learning. IEEE Internet of Things Journal, 9(10), 7110–7119. IEEE Internet of Things Journal.

Wang, Z., Xiao, J., Wang, L., & Yao, J. (2024). A novel federated learning approach with knowledge transfer for credit scoring. Decision Support Systems, 177, 114084.

Wei, J., Lowry, P. B., & Seedorf, S. (2015). The assimilation of RFID technology by Chinese companies: A technology diffusion perspective. Information & Management, 52(6), 628–642.

Wen, J., Zhang, Z., Lan, Y., Cui, Z., Cai, J., & Zhang, W. (2023). A survey on federated learning: Challenges and applications. International Journal of Machine Learning and Cybernetics, 14(2), 513–535.

Yoon, C., Lim, D., & Park, C. (2020). Factors affecting adoption of smart farms: The case of Korea. Computers in Human Behavior, 108, 106309.

Zhang, C., Xie, Y., Bai, H., Yu, B., Li, W., & Gao, Y. (2021). A survey on federated learning. Knowledge-Based Systems, 216, 106775.

Zhang, W., Lu, Q., Yu, Q., Li, Z., Liu, Y., Lo, S. K., Chen, S., Xu, X., & Zhu, L. (2021). Blockchain-Based Federated Learning for device failure detection in industrial IoT. IEEE Internet of Things Journal, 8(7), 5926–5937. IEEE Internet of Things Journal.

Framework Development of an ESG Course Series under System Dynamics: A Comparison of TKS and KSA Perspectives

Chiung-Chuan Lo, Ya-Ping Lin
Department of Business Administration, Shih Chien University. Taipei, Taiwan

Abstract

This study aims to develop an Environmental, Social, and Governance (ESG) course-series framework from the perspective of system dynamics and compare the differences between Task–Knowledge–Skills (TKS) and Knowledge–Skills–Abilities (KSA) in curriculum planning and learning outcome evaluation. As ESG has become an important topic in corporate governance, sustainable competitiveness, and management education, business education needs to move beyond one-way knowledge transmission toward a cumulative, practice-oriented, and assessable competence development model. This study adopts a literature-oriented conceptual modeling approach, integrating ESG education, competency models, task analysis, and system dynamics theory to develop causal loop diagrams (CLDs), stock-and-flow models, and a scenario simulation framework. The conceptual model indicates that the KSA perspective helps illustrate the long-term accumulation of ESG knowledge, analytical skills, and integrated decision-making ability. The TKS perspective helps translate ESG curriculum objectives into observable, measurable task outcomes. The conceptual scenario comparison suggests that the balanced hybrid course design best balances KSA competence growth, TKS task performance, and cognitive load control. The contribution of this research is to propose a dynamic governance framework for ESG courses, assisting business schools in curriculum design, learning assessment, and resource allocation.

Keywords: System dynamics, ESG education, TKS, KSA, curriculum design
Paper type: Academic Research Paper

1 Introduction

Environmental, Social, and Governance (ESG) has expanded from corporate sustainability disclosure and investment evaluation to a core competence area for organizational governance, strategic decision-making, and management education. Prior research shows that highly sustainability-oriented firms demonstrate distinctive organizational processes, stakeholder engagement, and long-term performance (Eccles et al., 2014), while ESG performance is also positively associated with corporate financial performance (Friede et al., 2015). Accordingly, business education should not treat ESG as a single knowledge topic, but should transform it into a cumulative, practice-oriented, and evaluable curriculum-based competence development system (Waddock & Lozano, 2013). Although ethics, corporate social responsibility (CSR), and sustainability have been gradually incorporated into global business school curricula, their integration still requires more systematic and multi-level curriculum design (Christensen et al., 2007; Setó-Pamies & Papaoikonomou, 2016).

Despite this development, three research gaps remain. First, existing ESG courses often emphasize sustainability concepts and case discussions, but pay less attention to how students progress from ESG knowledge to analytical skills, integrative abilities, and task performance. Second, learning evaluation frequently relies on static indicators, such as final examinations, questionnaires, or satisfaction surveys, which cannot adequately capture the dynamic relationships among learning engagement, project implementation, feedback, cognitive load, and competence accumulation. Third, although both Knowledge–Skills–Abilities (KSA) and Task–Knowledge–Skills (TKS) perspectives are relevant to curriculum design, their complementary roles in ESG course series have not been sufficiently examined. KSA emphasizes knowledge, skills, and abilities as the foundation of work and learning performance (Campion et al., 2011; Stevens & Campion, 1994), whereas TKS focuses on the connection among tasks, knowledge, and skills, helping translate competency requirements into actionable curriculum and performance standards (National Institute of Standards and Technology [NIST], 2021; Jonassen et al., 1998).

Based on these gaps, this study develops an ESG course-series framework from a system dynamics perspective and compares TKS and KSA in curriculum planning and learning outcome evaluation. Specifically, it clarifies the relationships among knowledge, skills, abilities, and task performance; constructs causal loop and stock-and-flow models of ESG learning outcomes; compares different curriculum design strategies; and proposes a dynamic decision-making framework for ESG curriculum governance. Methodologically, this study adopts a literature-oriented conceptual modelling approach. System dynamics is appropriate because it emphasizes feedback structures, time delays, nonlinear change, and policy simulation, which are useful for analyzing competence accumulation and learning outcomes in education systems (Sterman, 2000; Sweeney &

Sterman, 2007). Small system dynamics models can also support decision communication and scenario comparison (Ghaffarzadegan et al., 2011). This study contributes theoretically by integrating KSA, TKS, and system dynamics; methodologically by modelling ESG competence development; and practically by supporting ESG curriculum design, resource allocation, project-based learning, and competency-oriented assessment.

2 Literature review

2.1 ESG education and sustainability management competencies

ESG education has extended from corporate ethics, CSR, and sustainable development education to an important issue in business and management talent development. Existing studies have pointed out that management education must not only teach environmental, social, and governance concepts in response to sustainable transformation, but also cultivate students' systematic thinking, stakeholder analysis, strategic judgment, and problem-solving skills. Stubbs and Cocklin (2008) emphasize that sustainability education should assist business management students in changing their management mindset; Wiek et al. (2011) proposed that sustainability competencies should include systems thinking, anticipatory, normative, strategic, and interpersonal competencies. Therefore, ESG series courses should shift from a single knowledge transfer to a cross-curriculum, cross-task, and competency-oriented learning design.

2.2 KSA Competency Framework: Knowledge–Skills–Abilities

The KSA framework emphasizes the development process of learners from "knowledge" accumulation, "skill" proficiency to "ability" integration. Knowledge refers to conceptual, theoretical and institutional understanding; Skills refer to analytical, operational, and application-oriented proficiencies that can be developed through practice; Abilities refer to higher-level judgment, integration, and decision-making abilities. Campion et al. (2011) suggest that competency models can be used to define the knowledge, skills, abilities, and other attributes required for effective performance and can support training, assessment, and organizational development. When applied to ESG courses, KSA explains how students can move from ESG conceptual understanding to ESG data analysis skills, sustainability strategy integration, and governance decision-making abilities.

2.3 TKS Task-Based Competency Framework: Task–Knowledge–Skills

The TKS framework starts with a task, emphasizing first defining the specific tasks learners need to complete, and then deducing the knowledge and skills required to complete the task. This perspective is consistent with the task analysis and instructional design literature (NIST, 2021). Jonassen et al. (1998) pointed out that task analysis is an important foundation for instructional design, which can concretize learning objectives, work procedures, knowledge needs, and skill requirements. When applied to the ESG series, TKS can transform course outcomes into observable tasks such as ESG case diagnosis, sustainability report analysis, carbon emission data interpretation, stakeholder mapping, and ESG strategy proposals. This perspective helps to improve the operability and practical linkage of course assessment.

2.4 Comparison of KSA and TKS

The difference between KSA and TKS can be understood from two starting points of analysis: "ability formation" and "task performance". KSA is derived from competency modelling and job analysis research, emphasizing the knowledge, skills, and abilities required for individuals to perform effectively. In addition, abilities, skills, and task requirements can be systematically linked to explain how individual abilities and skills are translated into observable performance requirements (Fleishman et al., 1999). Campion et al. (2011) pointed out that competency models systematize the knowledge, skills, abilities, and other characteristics required for an organization or job, and support training, evaluation, and development decisions; Stevens and Campion (1994) also used team work as an example to illustrate that KSA can be used to identify individual-level competency requirements. Therefore, KSA is suitable for explaining how ESG learners can gradually develop ESG analysis skills, strategic integration skills, and sustainable decision-making skills from understanding sustainability concepts. The comparison between KSA and TKS is summarized in Table 1. This study therefore regards KSA and TKS as complementary competency frameworks rather than mutually exclusive classification tools.

Table 1. Comparison between KSA and TKS Perspectives

Dimension	KSA	TKS	ESG Course
Core logic	Focuses on the formation of learners' abilities, it emphasizes the gradual accumulation of knowledge, skills, and abilities	Focuses on task performance, first define the task, and then derive the required knowledge and skills	ESG courses need to focus on both ability development and task results
Analytical starting point	What knowledge, skills and abilities should students have?	What ESG tasks or projects should students complete?	Curriculum design can be developed in both directions based on competency goals and task requirements
Main aspects	Knowledge、Skills、Abilities	Task、Knowledge、Skills	KSA favors long-term competence development; TKS is biased towards short- and medium-term outcome verification
Knowledge definition	Basic knowledge of ESG concepts, sustainable governance, disclosure systems, stakeholder theory, etc	Expertise required to complete ESG tasks, such as ESG metrics, reporting guidelines, knowledge of carbon management	Introductory courses should establish a foundation of ESG knowledge and connect to task requirements
Skills definition	ESG data analysis, case interpretation, sustainability report analysis, problem-solving skills	Skills required to perform ESG tasks, such as data collection, metric analysis, presentations, and proposal skills	Intermediate courses can introduce data analysis, case discussions, and tool operations
Abilities / Task definition	Abilities refer to advanced integration abilities, such as strategic judgment, systems thinking, and sustainable decision-making	Tasks refer to specific observable tasks, such as ESG report analysis, carbon emission data interpretation, and sustainable strategy proposals	Advanced courses should combine project implementation with strategic integration training
Learning outcomes focus	competence accumulation, learning growth, and long-term development	Task completion, quality of results, and observable performance	Competency scales and task scoring standards can be designed at the same time
Assessment method	Pre- and post-tests, competence scales, learning process, reflection records, and ability growth curves	Project results, case reports, task performance, and performance rubrics	Mixed assessment should be adopted: competence assessment + task outcome evaluation
Advantages	It can show the phased development of ESG learners from knowledge to skills to advanced abilities	It can clearly link course activities, task requirements and performance outputs	It helps to balance the depth of the course with practical orientation
Limitations	If there is a lack of task scenarios, it may be difficult to correspond to specific course outcomes	If the design is overly task-oriented, it may overlook advanced abilities and long-term growth	Avoid focusing only on theory or task completion
Applicable course stage	It is suitable for beginner to advanced levels, especially suitable for competency maps and course sequence planning	Intermediate to advanced courses are more suitable and are especially suitable for project, practice, and outcome-oriented courses	The ESG series of courses should adopt the KSA-TKS dual-track structure
Value for ESG curriculum governance	Support long-term competence development, curriculum articulation, and competency map design	Support task design, outcome verification, and practice-oriented assessment	After integration, it can be used as a basis for ESG curriculum planning, evaluation, and resource allocation

2.5 System Dynamics and Learning Outcomes Assessment

System dynamics is suitable for analyzing educational issues with feedback, delay, and cumulative characteristics. Forrester (1961) and Sterman (2000) indicate that system dynamics describes complex system behavior through stocks, flows, and feedback loops. Learning outcomes are inherently cumulative, such as ESG

knowledge, analytical skills, decision-making ability, and task performance, which will change with learning engagement, practical experience, and feedback quality. Sweeney and Sterman (2007) also pointed out that systems thinking can help learners understand feedback, stock-flow, time delay, and nonlinear relationships. Ghaffarzadegan et al. (2011) emphasize that small system dynamic models can support policy communication and decision learning.

2.6 Literature gaps and research framework

Based on the above literature, there is still little research on ESG education that combines curriculum design, competency models, and dynamic simulations. Most studies emphasize the importance of sustainability competencies, but less explain how ESG knowledge, skills, abilities, and task performance accumulate, interact, and are moderated by constraints over time. In addition, although learning analytics research emphasizes data-based assessment and learning process tracking, it is still difficult to explain how different curriculum strategies affect long-term competence formation without causal models. Dawson et al. (2014) pointed out that the field of learning analysis has focused on understanding the learning process with data, but there is still a need for stronger theoretical and model linkages. Based on this, this study proposes a research framework of "ESG Curriculum Design–KSA/TKS Competence Formation–System Dynamics Simulation" to analyze the dynamic learning outcomes of ESG series courses using causal loop diagrams and stock-and-flow models. The research structure is summarized in Figure 1.

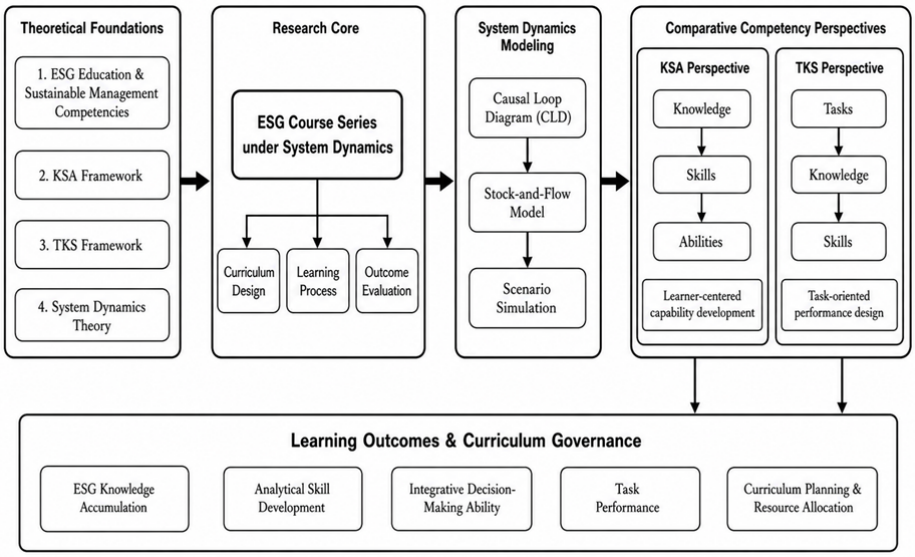


Figure 1: Research Framework of ESG Course Series under System Dynamics

3 Research methodology

3.1 Study design

This study adopts a conceptual system dynamics modeling approach, integrating ESG education, KSA competency model, TKS task analysis, and system dynamics literature to establish a dynamic analysis framework for ESG series courses. System dynamics is suitable for dealing with educational systems with feedback, delay, cumulative, and nonlinear characteristics, and can be used for scenario simulation through causal loop diagrams and stock-and-flow models (Forrester, 1961; Sterman, 2000). Sterman (2000) pointed out that system dynamics can be used as a modeling method for analyzing policy and strategic problems, and can be used to simulate complex systems and design effective policies.

The model-based results were derived through theoretical synthesis, variable mapping, causal structure identification, and conceptual scenario comparison rather than empirical parameter estimation.

3.2 Study scope and model boundaries

The boundaries of this research model are set in the ESG series of courses of the School of Business and Management, covering ESG fundamentals, sustainability report analysis, ESG data interpretation, stakeholder

analysis, case studies, and project implementation. The model does not address overall university governance or corporate ESG performance, but focuses on the process of forming students' knowledge, skills, abilities, and task performance in the curriculum.

3.3 Primary variable settings

The main stock includes ESG Knowledge Stock, ESG Analytical Skills, Integrative Decision-Making Ability, and ESG Task Performance. The main flows include Learning Engagement, Practice Intensity, Feedback Quality, and Ability Enhancement Rate. Auxiliary variables include Course Complexity, Cognitive Load, Learning Fatigue, Instructional Design Quality, and Task Difficulty. KSA variable design references competency model literature, considering knowledge, skills, and abilities as important foundations for effective performance (Campion et al., 2011); TKS, on the other hand, follows task analysis logic, deducing the required knowledge and skills from task requirements as the basis for course outcome design (Jonassen et al., 1998).

3.4 Causal loop model design

The causal loop model consists of four reinforcement loops and two equilibrium loops. R1 is the strengthening loop of "learning engagement - knowledge growth"; R2 is the "Project Implementation - Skill Growth" reinforcement loop; R3 is the "Feedback Correction - Task Performance" enhancement loop; R4 is the "Ability Integration–Decision Quality" reinforcement loop. B1 is the "cognitive load-learning fatigue" balance loop, and B2 is the "task pressure-decreased outcome quality" balance loop. This design echoes the emphasis on feedback, stock-flow, time delay, and nonlinear relationships in system dynamics (Sweeney & Sterman, 2007; Sterman, 2001).

3.5 Stock-and-flow model design

The stock-and-flow model is divided into two pathways: the KSA pathway, representing knowledge, skills, and abilities accumulation, and the TKS pathway, representing the translation of task requirements into required knowledge, required skills, and task performance. The KSA path is ESG Knowledge Stock -> ESG Analytical Skills -> Integrative Decision-Making Ability, emphasizing long-term competence accumulation. The TKS path is ESG Task Requirements -> Required Knowledge -> Required Skills -> Task Performance, emphasizing task performance and observable results. The two paths influence each other through feedback quality, implementation intensity, and cognitive load.

3.6 Scenario simulation design

Scenario simulation sets up six course strategies: S0 traditional teaching, S1 case-oriented, S2 project-oriented, S3 ESG data analysis training, S4 high-feedback intensity design, and S5 balanced hybrid course. Conceptual simulations are designed to compare the potential effects of each scenario on KSA competence growth, TKS task performance, and cognitive load. A small system dynamic model can assist in policy communication, scenario comparison, and leverage point identification, making it suitable as a decision support tool for ESG curriculum planning.

4 Research Results

4.1 ESG series curriculum structure planning results

Based on the sustainability education and curriculum design literature, this study proposes a five-stage ESG series curriculum framework of "Basic Cognition-Analysis Tools-Case Integration-Project Implementation-Reflection and Feedback". The first stage establishes ESG principles, sustainable governance, and disclosure system knowledge; the second stage introduces ESG indicators, data analysis, and sustainability performance evaluation tools; the third stage strengthens stakeholder analysis and systems thinking through corporate cases; the fourth stage develops ESG strategic integration ability through project-based application; and the fifth stage promotes reflection and revision through feedback from teachers, peers, and industry. These stages are consistent with experiential learning theory, which emphasizes learning through concrete experience, reflective observation, abstract conceptualization, and active experimentation (Kolb, 2014). This stage-based design echoes the idea that sustainability education should link competencies, pedagogies, and real-world problem-solving (Wiek et al., 2011; Lozano et al., 2017).

4.2 KSA vs. TKS Learning Outcomes Comparison Results

The KSA-based model suggests that ESG learning outcomes can be divided into three levels: knowledge accumulation, skill development, and ability integration, which are suitable for assessing students' long-term sustainable management abilities. The TKS-based model indicates that ESG course outcomes can be derived from task requirements by linking required knowledge and skills to observable task outcomes, such as ESG report analysis, carbon emission data interpretation, and sustainability strategy proposals. KSA emphasizes competence development, TKS emphasizes achievement verification; The integration of the two can prevent the competency model from remaining too abstract and task-oriented assessment from ignoring higher-order abilities. This dual-track design logic is supported by competency modeling and task analysis literature (Campion et al., 2011; Jonassen et al., 1998).

4.3 Causal loop analysis results

The causal loop model identifies four reinforcing loops and two balancing loops (Figure 2). R1 "Learning Engagement - Knowledge Growth" shows that learning engagement enhances ESG knowledge and further enhances learning confidence. R2 "Project Implementation - Skill Growth" shows that the intensity of implementation improves analytical skills and project quality. R3 "Feedback Correction - Task Performance" shows that high-quality feedback can improve student reflection and outcome correction. R4 "Ability Integration–Decision Quality" shows that knowledge and skill accumulation can improve ESG strategic judgment. In contrast, B1 "cognitive load - learning fatigue" and B2 "task pressure - decreased quality of results" indicate that excessive course intensity may inhibit learning effectiveness. This design is in line with the view of system dynamics to analyze complex systems with feedback, delay, and nonlinear relationships (Sterman, 2000; Sweeney & Sterman, 2007).

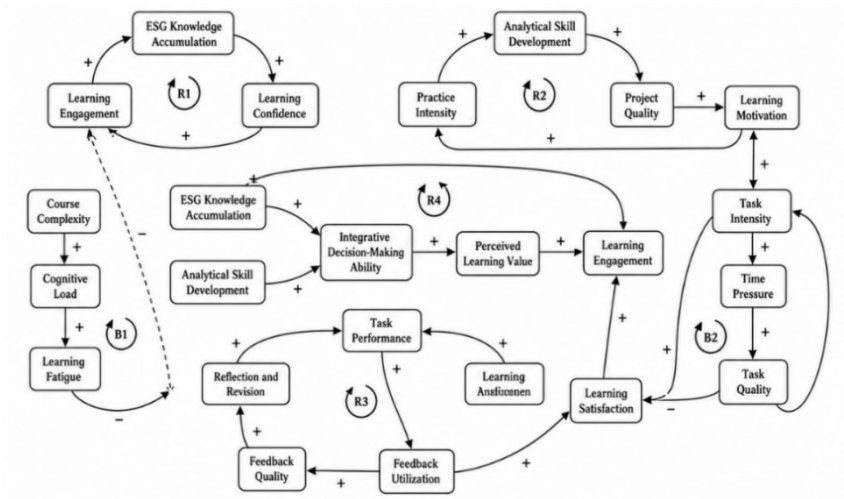


Figure 2: Causal Loop Diagram of ESG Course Series Learning System

4.3 Stock-and-flow model results

The stock-and-flow model (Figure 3) represents ESG Knowledge Stock, ESG Analytical Skills, Integrative Decision-Making Ability, and ESG Task Performance as the main stocks, while Learning Engagement, Practice Intensity, Feedback Quality, and Ability Enhancement Rate are specified as the main flows. The KSA path presents a cumulative process of "knowledge-skills-ability"; the TKS path presents a performance formation process of "task requirements-required knowledge-required skills-task performance". Both paths are moderated by cognitive load, course complexity, and feedback quality.

4.4 Scenario simulation results

The conceptual scenario comparison in Table 2 suggests that S0, the traditional lecture scenario, can steadily improve basic ESG knowledge but has limited effects on skills and task performance. Case-based learning enhances conceptual understanding and judgment, while project-based learning substantially improves TKS task performance but may increase cognitive load when feedback is insufficient. Data analytics training strengthens ESG analytical skills, and feedback-intensive design supports reflection, revision, and ability integration. Overall, the balanced hybrid course design provides the best balance among KSA competence growth, TKS task

outcomes, and manageable cognitive load. As a decision-support tool, the small system dynamics model helps compare curriculum strategies and support ESG curriculum governance (Ghaffarzadegan et al., 2011).

Table 2. Comparison of Scenario Simulation Outcomes

Scenario	KSA	TKS	Cognitive Load	Key simulation results
S0	Low to medium	Low	Low	It can steadily improve basic ESG knowledge but has limited effects on analytical skills, integrative decision-making ability, and task outcomes
S1	Medium	Medium	Low to medium	It helps students understand the context of ESG issues, stakeholder relationships, and corporate sustainability decision-making scenarios
S2	Medium to high	High	High	It improves ESG task completion and project quality but may increase time pressure and learning fatigue.
S3	High	Medium to high	Medium to high	It can strengthen ESG indicator interpretation, sustainability report analysis, and data processing skills, thereby improving KSA skill stock
S4	High	High	Medium	Through feedback from teachers, peers, and the industry, it can promote reflection and correction, task quality improvement, and integrative ability development
S5	Highest	High to very high	Medium	At the same time, it combines cases, projects, data analysis, and feedback mechanisms to balance long-term competence growth, task performance, and manageable cognitive load

5 Discussion

5.1 Theoretical implications

This study regards the ESG series of courses as a dynamic system that links competence accumulation and task performance, making up for the shortcomings of past ESG education research that focuses on course content and single-point learning assessment. The ESG and sustainability education literature points out that ESG and sustainability management education should cultivate systems thinking, strategic skills, collaboration skills, and problem-solving skills, rather than just imparting knowledge (Wiek et al., 2011; Lozano et al., 2017). Therefore, this study uses system dynamics to link competency models and curriculum design to deepen the theoretical explanatory power of ESG education.

5.2 Complementarity of KSA and TKS

KSA emphasizes the phased development of knowledge, skills, and abilities, which is suitable for explaining how ESG learners move from conceptual understanding to analysis and integrated decision-making. TKS is task-oriented, suitable for transforming ESG course outcomes into observable and evaluable task performance. The competency model literature supports the use of KSA for training, assessment, and development decision-making (Campion et al., 2011), while the task analysis literature emphasizes that instructional design needs to be developed by tasks, knowledge, and skill requirements (Jonassen et al., 1998). The integration of the two can support both competence development and outcome verification (Table 3).

Table 3: ESG Course Series Architecture under KSA and TKS

Course modules	KSA	TKS	Main activities	Learning assessment
ESG Foundation	Build ESG basic knowledge	Understand task-related ESG concepts	Lectures, readings, short case discussions	Concept test, participation, learning notes
ESG Governance and Disclosure	Develop knowledge of ESG governance and disclosure	Define sustainability report analysis tasks	Report review, materiality analysis, ESG case study	Report summary, materiality issue table
ESG Data and Analytical Tools	Strengthen ESG analytical skills	Apply required skills to data-based tasks	Indicator analysis, data interpretation, visualization	ESG data analysis report, tool-use task
ESG Case Integration	Integrate knowledge and skills into judgment	Diagnose ESG problems through cases	Case discussion, stakeholder mapping, group debate	Case diagnosis report, oral presentation

Course modules	KSA	TKS	Main activities	Learning assessment
ESG Project-Based Application	Develop integrative decision-making and problem-solving abilities	Complete ESG task outcomes	Project work, strategy proposal, industry feedback	ESG project report, proposal presentation
Reflection and Feedback	Enhance higher-order integrative abilities through revision	Improve task performance through feedback	Reflection notes, peer review, teacher feedback	Revised output, reflection report, peer assessment
Curriculum Governance and Learning Analytics	Monitor KSA competence accumulation	Track TKS task performance	Pre/post-test, portfolio review, learning analytics	KSA competence indicators, TKS task performance indicators

5.3 Curriculum governance implications

This study shows that ESG course governance should not be based solely on a single course, a single test, or satisfaction, but should be evaluated based on course series, learning journey, task outcomes, and feedback mechanisms. System dynamics emphasizes feedback, delay, and stock accumulation, which can help course designers identify the dynamic relationship between learning engagement, project implementation, cognitive load, and competence growth (Sterman, 2000; Sweeney & Sterman, 2007). Therefore, curriculum governance should focus on curriculum load allocation, feedback density, and cross-stage competence alignment (Ghaffarzadegan et al., 2011).

5.4 Practical implications

In practice, this study can assist business schools in planning ESG course maps: beginner courses to build ESG knowledge, intermediate courses to strengthen data analysis and case judgment, and advanced courses to cultivate strategic integration skills through project implementation. Small system dynamic models can also be used as a communication tool for curriculum decision-making to compare the impact of different programs such as case teaching, project orientation, data analysis training, and high-feedback design on learning outcomes (Ghaffarzadegan et al., 2011).

6 Conclusion

6.1 Research conclusions

This study constructs the ESG series curriculum structure from the perspective of system dynamics and compares the two competency perspectives of KSA and TKS. The model indicates that KSA can illustrate the long-term accumulation of ESG knowledge, skills, and higher-order abilities. TKS transforms course objectives into observable, evaluable task outcomes. Competency model research emphasizes that knowledge, skills, and abilities can be used as the basis for training, evaluation, and development decision-making (Campion et al., 2011). Sustainability education research also indicates that ESG courses should cultivate systems thinking, strategic skills, and problem-solving skills, rather than just imparting knowledge (Wiek et al., 2011). Therefore, the integration of KSA and TKS helps to balance “competence development” and “outcome validation”, echoing the dynamic curriculum governance concept proposed in the original abstract of this study.

6.2 Study limitations

This study is a conceptual model construction, and the model parameter calibration and validity test have not yet been conducted using actual student learning data. Secondly, the effectiveness of ESG courses may be affected by student background, teachers' teaching competence, curriculum resources, industry connections, and school system support, so the model results still need to be adjusted according to different educational scenarios. Furthermore, although the system dynamic model can present feedback structure and long-term trends, if there is no empirical data support, the scenario simulation results should still be regarded as decision-making aids rather than final predictions.

6.3 Future research directions

Future research can collect student learning journeys, project outcomes, feedback records, and pre- and post-test data to calibrate the stock-and-flow model and test the explanatory power of KSA and TKS indicators on ESG learning outcomes. Secondly, different course designs, such as case teaching, project-oriented, data

analysis training, and high feedback mechanisms, can be compared to the differences in long-term competence accumulation and task performance. System dynamics studies have pointed out that small models are helpful for policy communication, scenario comparison, and leverage point identification (Ghaffarzadegan et al., 2011), so ESG course governance dashboards can also be developed in the future to support course adjustment and learning early warning.

7 Declaration of Ethics

This study does not address research ethics issues.

8 Artificial Intelligence Declaration

Artificial intelligence tools were used to assist in language refinement, preliminary literature organization, and structural drafting. The authors reviewed, verified, revised, and take full responsibility for all arguments, citations, interpretations, and final content.

References

Campion, M. A., Fink, A. A., Ruggeberg, B. J., Carr, L., Phillips, G. M., & Odman, R. B. (2011). Doing competencies well: Best practices in competency modeling. *Personnel Psychology*, *64*(1), 225–262.

Christensen, L. J., Peirce, E., Hartman, L. P., Hoffman, W. M., & Carrier, J. (2007). Ethics, CSR, and sustainability education in the Financial Times top 50 global business schools: Baseline data and future research directions. *Journal of Business Ethics*, *73*(4), 347–368.

Dawson, S., Gašević, D., Siemens, G., & Joksimovic, S. (2014, March). Current state and future trends: A citation network analysis of the learning analytics field. In Proceedings of the fourth international conference on learning analytics and knowledge (pp. 231–240).

Eccles, R. G., Ioannou, I., & Serafeim, G. (2014). The impact of corporate sustainability on organizational processes and performance. *Management Science*, *60*(11), 2835–2857.

Fleishman, E. A., Costanza, D. P., & Marshall-Mies, J. C. (1999). Abilities, skills, and task requirements. In D. R. Ilgen & E. D. Pulakos (Eds.), *The changing nature of performance: Implications for staffing, motivation, and development* (pp. 175–211). Jossey-Bass.

Forrester, J. W. (1961). *Industrial dynamics.* MIT Press.

Friede, G., Busch, T., & Bassen, A. (2015). ESG and financial performance: aggregated evidence from more than 2000 empirical studies. *Journal of Sustainable Finance & Investment*, *5*(4), 210–233.

Ghaffarzadegan, N., Lyneis, J., & Richardson, G. P. (2011). How small system dynamics models can help the public policy process. *System Dynamics Review*, *27*(1), 22–44.

Jonassen, D. H., Tessmer, M., & Hannum, W. H. (1998). *Task analysis methods for instructional design*. Routledge.

Kolb, D. A. (2014). *Experiential learning: Experience as the source of learning and development*. FT Press.

Lozano, R., Merrill, M. Y., Sammalisto, K., Ceulemans, K., & Lozano, F. J. (2017). Connecting competences and pedagogical approaches for sustainable development in higher education: A literature review and framework proposal. *Sustainability*, *9*(10), 1889.

National Institute of Standards and Technology. (2021). Task, Knowledge, Skill (TKS) statements authoring guide for workforce frameworks. https://www.nist.gov/system/files/documents/2021/07/30/TKS_Authoring_Guide13apr2021-508Compliant.pdf

Setó-Pamies, D., & Papaoikonomou, E. (2016). A multi-level perspective for the integration of ethics, corporate social responsibility and sustainability (ECSRS) in management education. *Journal of Business Ethics*, *136*(3), 523–538.

Sterman, J. D. (2000). *Business dynamics: Systems thinking and modeling for a complex world*. McGraw-Hill.

Sterman, J. D. (2001). System dynamics modeling: tools for learning in a complex world. *California management review*, 43(4), 8–25.

Stevens, M. J., & Campion, M. A. (1994). The knowledge, skill, and ability requirements for teamwork: Implications for human resource management. *Journal of management*, *20*(2), 503–530.

Stubbs, W., & Cocklin, C. (2008). Teaching sustainability to business students: shifting mindsets. International *Journal of Sustainability in Higher Education*, *9*(3), 206–221.

Sweeney, L. B., & Sterman, J. D. (2007). Thinking about systems: Student and teacher conceptions of natural and social systems. System Dynamics Review: *The Journal of the System Dynamics Society*, *23*(2-3), 285–311.

Waddock, S., & Lozano, J. M. (2013). Developing more holistic management education: Lessons learned from two programs. *Academy of Management Learning & Education*, *12(2)*, 265–284.

Wiek, A., Withycombe, L., & Redman, C. L. (2011). Key competencies in sustainability: a reference framework for academic program development. *Sustainability science*, *6(2)*, 203–218.

A System Dynamics Analysis of AI-Empowers Smart Aquaculture Technologies and Their ESG Performance Impacts

Ju-We Chen[1], Chi-Yuan Lin[2], Tsu-Erh Kuo[1]
[1] Department of Business Administration, Shih Chien University. Taipei, Taiwan
[2] Fisheries Research Institute. New Taipei City, Taiwan

Abstract

This study aims to explore the dynamic impact of Artificial Intelligence (AI)-empowered smart aquaculture technologies on environmental, social, and governance (ESG) performance. Faced with challenges in the aquaculture industry such as water quality fluctuations, overfeeding, sediment accumulation, energy consumption, and governance transparency, this study uses system dynamics (SD) as a method to construct an analysis framework of "data sensing-AI decision-making-operational control-ESG performance", focusing on three core technologies: long-term dissolved oxygen (DO) sensing, AI-empowered precision feeding, and intelligent sludge removal. The study analyzes the feedback relationship between water quality stability, fish biomass, sediment accumulation, farmer acceptance, data traceability, and ESG performance through causal loop diagrams and stock-and-flow models. The results show that AI data learning can improve decision-making quality, precision feeding can improve feed utilization efficiency and fish growth, intelligent sludge removal can reduce environmental risks, and increased automation may also bring energy load and carbon emission pressure. The scenario comparison results suggest that, compared with conventional aquaculture, a full smart aquaculture system integrating DO sensing, AI-empowered precision feeding, intelligent sludge removal, and ESG governance dashboards can generate cumulative environmental improvements, reduce labor risks, and enhance governance transparency. The main contribution of this study is to conceptualize AI-empowered smart aquaculture technologies as knowledge-intensive assets that transform sensing data, algorithmic inference, and expert rules into operational decisions and auditable ESG evidence.

Keywords: System Dynamics; Artificial Intelligence of Things (AIoT); Smart Aquaculture; ESG Performance; Precision Feeding

Paper type：Academic Research Paper

1 Introduction

The global aquaculture industry is facing multiple challenges such as improving production efficiency, environmental sustainability pressures, and governance transparency requirements. With the expansion of intensive aquaculture models, issues such as water quality fluctuations, dissolved oxygen instability, overfeeding, sediment accumulation, energy consumption, and disease risks have become important factors affecting aquaculture performance and sustainable development. Garlock et al. pointed out that the sustainability of aquaculture should not be measured solely by yield or economic benefits, but should also include comprehensive performance evaluations from environmental, economic, and social aspects. Its Aquaculture Performance Indicators also show significant differences in environmental impact, social benefits, and economic resilience between different aquaculture systems (Garlock et al., 2024). In addition, feed efficiency and the use of resources such as fishmeal and fish oil have a long-term impact on the environmental footprint of aquaculture, indicating that precision feeding and resource optimization have become key issues in the sustainable transformation of the industry (Tacon & Metian, 2008).

In recent years, artificial intelligence, Internet of Things, and sensing technologies have been rapidly introduced into aquaculture farms, gradually shifting smart aquaculture from traditional empirical management to real-time monitoring, data-driven decision-making, and automated control. Artificial Intelligence of Things (AIoT) technologies can collect real-time data such as dissolved oxygen, temperature, pH, salinity, and fish behavior through sensors, and use Artificial Intelligence (AI) models for water quality prediction, disease monitoring, and feeding decision support (Huang & Khabusi, 2025). Among them, dissolved oxygen monitoring is regarded as a core technology for maintaining fish health and water quality stability, and online intelligent dissolved oxygen detection can improve the real-time and management accuracy of environmental monitoring (Wei et al., 2019). At the same time, computer vision and fish feeding behavior recognition technologies are gradually being applied to intelligent feeding to help determine the feeding status of fish, reduce residual bait, and improve feed utilization (Xiao et al., 2025).

However, existing studies have focused on the accuracy, equipment performance, or short-term operational improvement of a single technology, and there is still a lack of integrated analysis of how AI-empowered

technologies can further impact environmental, social, and governance performance through feedback structures, time delays, and system interactions. Aquaculture is essentially a highly interdependent and complex system, and the amount of feeding affects water quality and sediment accumulation, while water quality stability affects fish growth, disease risk, and energy use. If technology performance is evaluated only based on static indicators, it is easy to overlook the cumulative and delayed nature of environmental, social, and governance (ESG) benefits. Muralidhar et al. (2023) noted that dynamic models help understand the complex interplay between environmental factors, growth, survival, and disease risk in aquaculture; Recent studies on recirculating aquaculture systems have also shown that system dynamics can be used for water quality management, resource use, and sustainability efficiency analysis (Muralidhar et al., 2023; Udayakumar et al., 2025).

Based on the above awareness, this study aims to construct a system dynamics analysis framework to explore how AI-empowered smart aquaculture technologies affect ESG performance. This study focuses on three verified core technologies: long-term dissolved oxygen sensing module, AI-empowered precision feeding system, and intelligent sludge removal technology and analyze their dynamic impacts on water quality stability, feed efficiency, sediment risk, labor load, data traceability, and governance transparency through causal loop diagrams and stock-and-flow models. The contributions of this research are threefold. First, it applies system dynamics to the analysis of AI-empowered smart aquaculture and ESG performance. Second, it transforms field-verified technical parameters into a basis for long-term sustainability-oriented scenario comparison. Third, it responds to IFKAD's focus on knowledge assets and sustainable organizations by conceptualizing AI technologies as knowledge-intensive assets embedded with sensing data, expert logic, and governance knowledge. This view also echoes the argument of Eccles et al. on corporate sustainability governance, that sustainable performance needs to be achieved through organizational processes, governance mechanisms, and long-term value creation (Eccles et al., 2014).

2 Literature Review

Artificial Intelligence of Things (AIoT) is an intelligent system that combines sensors, edge computing, data transmission, and artificial intelligence algorithms, and its core value is to transform field data into actionable decision-making knowledge in real time. In aquaculture scenarios, AIoT can continuously collect data such as dissolved oxygen, temperature, pH, salinity, water flow, fish behavior, and equipment status, and support water quality management, disease early warning, feeding judgment, and automated control through machine learning, computer vision, and predictive models. Huang and Khabusi (2025) pointed out that the main applications of AIoT in aquaculture include real-time environmental monitoring, data-driven decision-making, and automated operations, which can improve farming efficiency, sustainability, and production stability. Computer vision studies have also shown that fish monitoring can cover aspects such as fish body identification, biomass estimation, behavioral classification, and health analysis, indicating that smart aquaculture is moving from single-point monitoring to integrated intelligent management (Al-Abri et al., 2025).

The technical foundation of smart aquaculture is mainly based on three types of systems: environmental sensing, precision feeding, and automated maintenance. Dissolved oxygen is a key water quality factor affecting fish metabolism, feeding, growth, and survival, and online intelligent dissolved oxygen detection technologies can enhance real-time monitoring and dynamic management capabilities (Wei et al., 2019). In terms of feeding management, overfeeding not only causes feed waste but also increases the accumulation of residual bait, ammonia nitrogen, and sediment, thereby deteriorating water quality and increasing the risk of disease. Therefore, AI-empowered precision feeding systems can adjust the timing and amount of feeding through fish feeding behavior recognition, reducing environmental impact and improving feed utilization efficiency. Xiao et al. (2025) pointed out that computer vision has been widely used in the quantification and identification of fish feeding behavior, providing an important technical foundation for intelligent feeding. The long-term dissolved oxygen sensing, AI-empowered precision feeding, and intelligent sludge removal technologies focused in this article are the combination of core technologies that integrates environmental data, behavior interpretation, and automated maintenance into the aquaculture decision-making system.

System Dynamics (SD) provides an important method for analyzing the complex causal relationships of intelligent farming. Forrester (1961) proposed that system behavior often stems from the interaction between stocks, flows, feedback loops, and time delays; Sterman (2000) further pointed out that SD is suitable for dealing with nonlinearity, delay effects, and policy simulation problems. Aquaculture systems have typical dynamic complexities, such as feeding volume affecting fish growth and water quality load, while changes in water quality can feed back on feeding, disease risk, energy use, and maintenance needs. Muralidhar et al. (2023) pointed out that dynamic models help understand the interaction of environmental variables, biological growth, and

management interventions in coastal aquaculture. Udayakumar et al. (2025) also applied SD to water quality management in recirculating aquaculture systems, emphasizing its ability to support resource efficiency and sustainable operational analysis.

ESG governance is an important theoretical basis for evaluating the value of smart aquaculture technologies. Traditional farming performance mostly focuses on yield, cost, and income, but sustainable farming needs to incorporate environmental, social, and governance aspects at the same time. Garlock et al. (2024) analyzed 57 aquaculture systems and 88 indicators using Aquaculture Performance Indicators, pointing out that aquaculture sustainability needs to measure environmental, economic, and social outcomes at the same time, rather than just based on a single production efficiency. In this context, the environmental aspects can include water quality stability, nutrient load, sediment accumulation, energy use, and carbon emissions. The social aspect can cover labor intensity, occupational risks, farmers' acceptance and knowledge transfer; The governance aspect involves data traceability, auditability of key performance indicators, transparency in decision-making, and policy consistency. Corporate sustainability research also points out that highly sustainable organizations usually have more comprehensive governance processes, stakeholder communication, and long-term performance orientation, indicating that ESG is not an external disclosure activity, but rather a governance capability embedded in organizational processes and decision-making systems (Eccles et al., 2014).

AIoT and smart aquaculture literature explains technical feasibility, SD literature provides methods for analyzing feedback structures and delay effects, and ESG governance literature provides a sustainable performance evaluation framework. However, existing studies have mostly discussed individual technical performance, aquaculture management, or sustainability indicators, and rarely regarded AI-empowered smart aquaculture technologies as a knowledge asset embedded with sensor data, expert knowledge, and governance logic, and further analyzed its long-term dynamic impact on ESG performance. Therefore, this paper attempts to integrate AIoT, smart aquaculture, system dynamics, and ESG governance perspectives to construct an analytical framework that can explain how smart aquaculture technology generate sustainable performance through feedback loops, stock-and-flow structures, and model-based scenario analysis.

3 Research methodology

3.1 Model boundary and variables

This study adopts system dynamics (SD) to construct a dynamic analytical model of AI-empowered smart aquaculture and its impact on environmental, social, and governance (ESG) performance. SD explains system behavior through the interactions among stocks, flows, feedback loops, and time delays, and is suitable for analyzing nonlinear relationships, delayed effects, and policy scenarios (Forrester, 1961; Sterman, 2000). Following Barlas (1996), model credibility was considered through structural verification, dimensional consistency checks, extreme-condition tests, and sensitivity analysis.

The model boundary is defined as the AI-empowered smart aquaculture operation system, covering five subsystems: water quality management, feeding and fish biomass growth, sludge accumulation and environmental risk, energy and labor input, and data governance and technology adoption. The model focuses on three core technologies: long-term dissolved oxygen (DO) sensing, AI-empowered precision feeding, and intelligent sludge removal. These technologies are treated as knowledge-intensive assets that transform sensing data, algorithmic inference, and expert rules into operational decisions and auditable ESG evidence. The major stocks include water quality stability, fish biomass, sludge accumulation, farmer acceptance, data traceability, and ESG performance. The major flows include feed input rate, growth rate, mortality rate, sludge removal rate, daily energy consumption, labor input, and data accumulation rate.

Table 1. Key Variables and Model Structure

Variable	Type	ESG	Description
Water Quality Stability	Stock / Auxiliary	E	Stability of DO, pH, and water conditions
Feed Input Rate	Flow	E	Daily quantity of feed applied
Fish Biomass	Stock	E / Economic	Accumulated fish biomass and survival outcome
Sludge Accumulation	Stock	E	Sediment, residual feed, and waste accumulation
Daily Energy Consumption	Flow	E	Daily energy used by sensors, aeration, and devices
Labor Intensity	Auxiliary	S	Manual workload and maintenance burden
Farmer Acceptance	Stock	S / G	Adoption willingness and trust

Variable	Type	ESG	Description
Data Traceability	Stock	G	Accumulated auditable production and environmental data
ESG Performance	Composite Indicator	E / S / G	Integrated sustainability performance outcome

3.2 Data sources and parameter setting

The data sources are divided into three categories. First, the model incorporates field-verified parameters from the Ministry of Agriculture project, "Innovative Aquaculture Long-Term Sensing Chip and Verification of Intelligent Pollutant Discharge." The key parameters include long-term DO sensing stability exceeding seven consecutive days, AI-empowered precision feeding decision accuracy above 95%, and intelligent sludge removal coverage above 90%. Second, smart aquaculture and AIoT literature is used to support the technical logic of DO sensing, fish feeding behavior recognition, and data-driven aquaculture management. Online DO detection supports real-time water quality monitoring and risk control (Wei et al., 2019), while computer vision-based feeding behavior recognition supports AI-empowered precision feeding (Xiao et al., 2025). AIoT studies also show that real-time sensing, predictive analytics, and automated control can improve aquaculture efficiency and sustainability (Huang & Khabusi, 2025). Third, ESG and sustainable aquaculture studies are used to transform environmental, social, and governance dimensions into measurable indicators. Aquaculture sustainability should be evaluated through integrated environmental, economic, and social indicators rather than production output alone (Garlock et al., 2024).

The simulation horizon was set to 120 days, with one day as the basic time unit. Parameters were derived from field validation records, literature-based ranges, and expert judgment.

3.3 Stock-and-Flow Equations

To operationalize the stock-and-flow structure, five core equations were specified. In the following equations, t denotes the current time step and dt denotes the simulation time interval.

· Equation 1. Fish Biomass Accumulation:

$$Fish_Biomass(t)=Fish_Biomass(t-dt)+(Growth_Rate(t)-Mortality_Rate(t))\times dt$$

Fish biomass increases through growth rate and decreases through mortality rate. AI-empowered precision feeding improves feed utilization and water quality, thereby increasing growth and reducing mortality risk.

· Equation 2. Sludge Accumulation:

$$Sludge_Accumulation(t)=Sludge_Accumulation(t-dt)+(Residual_Feed_Generation(t)-Sludge_Removal_Rate(t))\times dt$$

Sludge accumulation increases with residual feed generation and decreases through intelligent sludge removal. Overfeeding increases sediment and nutrient load, while sludge removal reduces environmental risk.

· Equation 3. Water Quality Stability:

$$Water_Quality_Stability(t)=0.35DO_Stability(t)+0.25Feeding_Efficiency(t)+0.25Sludge_Reduction(t)+0.15Nutrient_Load_Reduction(t)$$

Water quality stability is determined by DO stability, feeding efficiency, sludge reduction, and nutrient load reduction.

· Equation 4. Data Traceability Accumulation:

$$Data_Traceability(t)=Data_Traceability(t-dt)+Data_Accumulation_Rate(t)\times dt$$

$$Data_Accumulation_Rate(t)=Sensor_Coverage(t)\times Data_Quality(t)\times Governance_Audit_Intensity(t)$$

Data traceability accumulates through sensor coverage, data quality, and governance audit intensity, supporting decision auditability and ESG reporting transparency.

· Equation 5. Farmer Acceptance Dynamics:

$$Farmer_Acceptance(t)=Farmer_Acceptance(t-dt)+(Trust_Building_Rate(t)-Resistance_Rate(t))\times dt$$

$$Trust_Building_Rate(t)=Decision_Accuracy(t)\times Operational_Benefit(t)\times Training_Effectiveness(t)$$

Farmer acceptance increases when AI decisions are accurate, operational benefits are observable, and training is effective; it decreases when perceived complexity or maintenance burden increases.

3.4 ESG performance calculation

ESG performance was operationalized as a composite indicator integrating environmental, social, and governance dimensions. Because the indicators are measured in different units, all variables were normalized to a 0–100 scale before aggregation. Positive indicators, such as water quality stability, farmer acceptance, and

data traceability, were normalized using min–max transformation. Negative indicators, such as sludge accumulation, energy consumption, labor intensity, and carbon emissions, were reverse-normalized.

The overall ESG performance score was calculated as:

$$ESG_Performance(t)=0.40EPI(t)+0.30SPI(t)+0.30GPI(t)$$

The environmental dimension was assigned a slightly higher weight because aquaculture sustainability is directly constrained by water quality, feeding efficiency, sludge accumulation, and energy use. The three sub-indices are:

$$EPI(t)=0.30WQS(t)+0.25FE(t)+0.20SLR(t)+0.15ER(t)+0.10CR(t)$$

$$SPI(t)=0.30LR(t)+0.25ORR(t)+0.25FA(t)+0.20KT(t)$$

$$GPI(t)=0.35DT(t)+0.25DA(t)+0.25KPI(t)+0.15PA(t)$$

where WQS denotes water quality stability, FE feeding efficiency, SLR sludge load reduction, ER energy reduction, CR carbon reduction, LR labor reduction, ORR occupational risk reduction, FA farmer acceptance, KT knowledge transfer, DT data traceability, DA decision auditability, KPI refers to key performance indicator transparency, and PA refers to policy alignment.

3.5 Scenario comparison design

The scenario comparison design includes four technology adoption pathways. S0 represents conventional aquaculture without AI-enabled smart technologies. S1 introduces long-term DO sensing to improve water quality monitoring and reduce environmental uncertainty. S2 further adds AI-empowered precision feeding to improve feed utilization, reduce residual feed, and support fish biomass growth. S3 represents the full smart aquaculture system integrating DO sensing, AI-empowered precision feeding, intelligent sludge removal, and ESG governance dashboards.

By comparing water quality stability, feed efficiency, sludge accumulation, labor intensity, farmer acceptance, data traceability, and ESG performance across the four scenarios, this study examines whether AI-empowered smart aquaculture technologies generate cumulative ESG improvements. The scenario comparison also identifies the reinforcing and balancing mechanisms through which smart aquaculture technologies transform technical performance into long-term sustainability outcomes.

4 Research Results

Based on the system dynamics model structure and scenario comparison, this study summarizes the AI-empowered smart aquaculture system into a dynamic structure of "data sensing-AI decision-making-operational control-ESG performance". Causal loop analysis shows that long-term dissolved oxygen sensing, AI-empowered precision feeding, and intelligent sludge removal are not individual equipment benefits, but jointly affect water quality stability, fish growth, environmental load, labor demand, and governance transparency through feedback mechanisms. This result echoes the view of the system dynamics literature on complex aquaculture systems, i.e., that water quality, feeding, biological metabolism, and management interventions are highly interactive, and their nonlinearity and delay effects need to be grasped through dynamic models (Muralidhar et al., 2023; Udayakumar et al., 2025).

The causal loop diagram presents two main reinforcing loops (Figure 1): R1 and R2. R1's "AI Data Learning - Decision Quality" shows that the accumulation of sensing data can improve the learning quality of AI models, thereby improving dissolved oxygen monitoring, feeding judgment, and intelligent sludge removal. After the quality of decision-making is improved, the operational data is more complete, further strengthening the AI judgment ability. AIoT research also points out that smart aquaculture can improve production stability through real-time environmental monitoring, data-driven decision-making, and automated control (Huang & Khabusi, 2025). R2's "Precision Feeding - Fish Growth" shows that improving the accuracy of AI feeding can reduce residual bait and feed waste, improve water quality, and increase fish growth and survival rates. Research on fish feeding behavior recognition also supports the use of computer vision as a technical basis for intelligent feeding decisions (Xiao et al., 2025).

The model also identifies three balancing loops: B1, B2, and B3. Balancing loop B1, “Overfeeding–Water Quality Deterioration,” shows that when feed input exceeds fish demand, residual feed and nutrient load increase, leading to lower dissolved oxygen, sediment accumulation, higher disease risk, and ultimately reduced growth performance. Balancing loop B2, “Intelligent Sludge Removal–Environmental Risk Control,” shows that sediment accumulation increases environmental risks and maintenance needs, while intelligent sludge removal systems reduce sediment load and delay water quality deterioration. Balancing loop B3, “Automation–Energy Load,” indicates that higher automation may increase electricity consumption, operational costs, and carbon emission pressure, thereby requiring energy-saving management mechanisms.

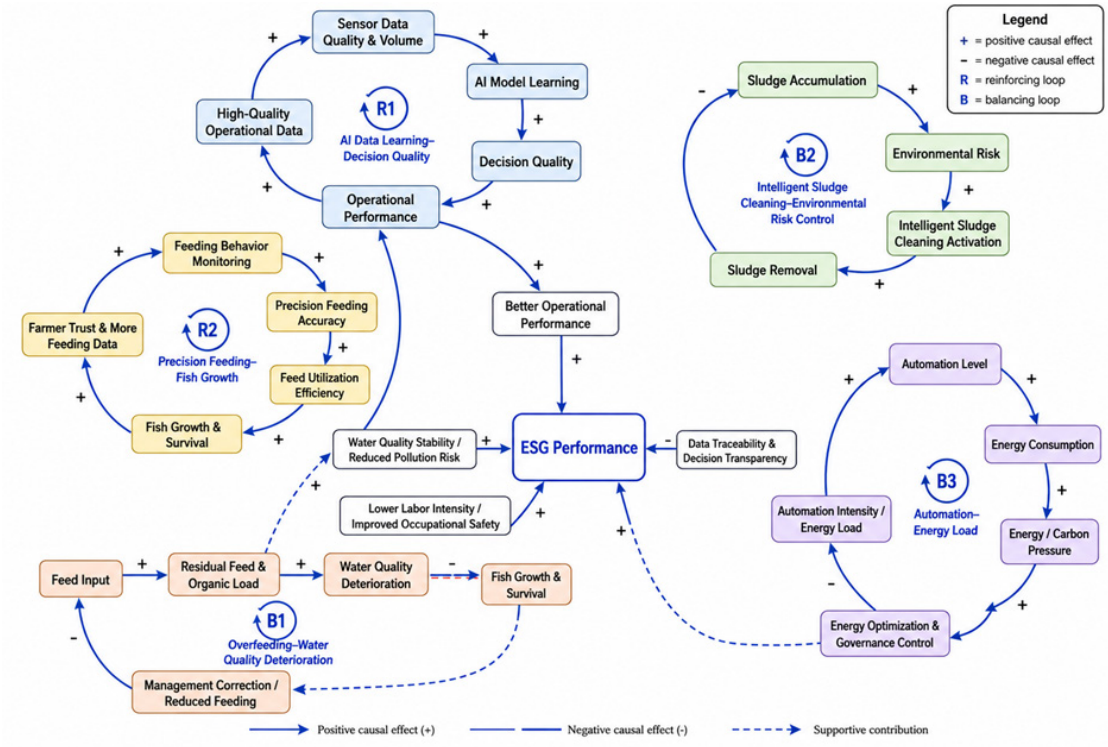


Figure 1. Causal Loop Diagram of the AI-Empowered Smart Aquaculture ESG System

In terms of stock-and-flow structure (Figure 2), this study sets "water quality stability", "fish biomass", "sediment accumulation", "farmer acceptance", "data traceability" and "ESG performance" as the main stocks. The model sets feed input rate, fish growth rate, mortality rate, intelligent sludge removal efficiency, energy consumption rate, labor input rate, and data accumulation rate as the main flows. The technical parameters of this study, including long-term dissolved oxygen sensing stability for more than seven days, AI-empowered precision feeding decision accuracy of more than 95%, and intelligent sludge removal coverage rate of more than 90%, can be used as the basis for model parameterization and scenario comparison.

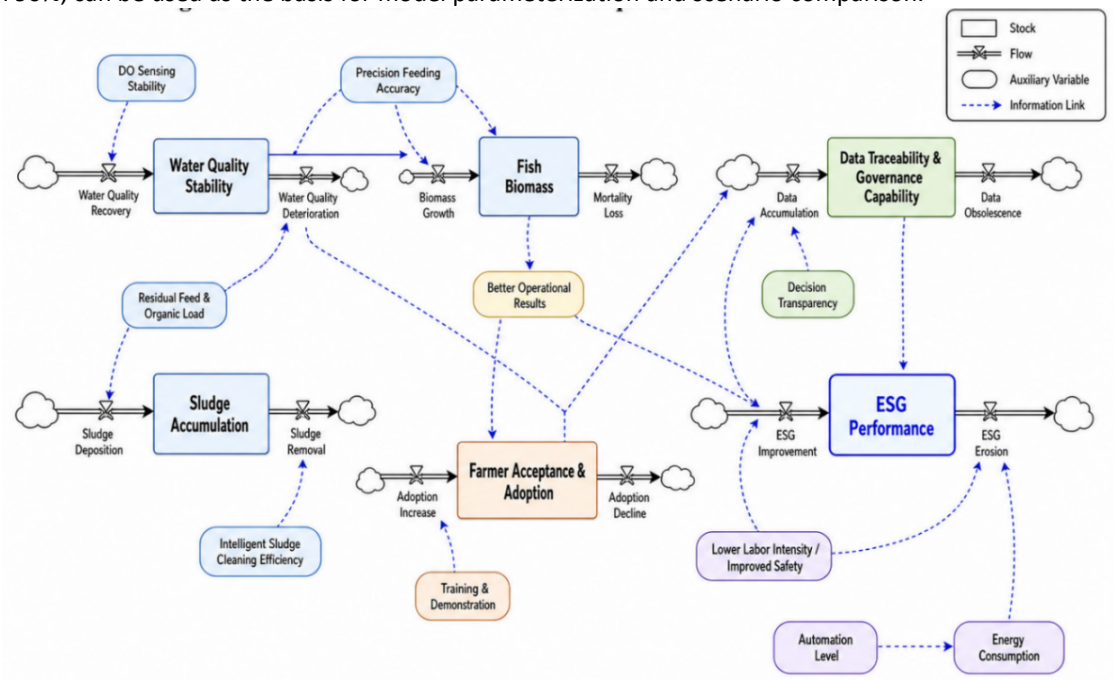


Figure 2. Stock-and-Flow Model of Smart Aquaculture ESG Performance

The model-based scenario comparison (Table 2) suggests that, under the S0 conventional aquaculture scenario, water quality fluctuations, sediment accumulation, and labor load remain relatively high, while ESG

data traceability is limited. In S1, the introduction of long-term DO sensing reduces water quality uncertainty and improves environmental monitoring capability. In S2, the addition of AI-empowered precision feeding improves feed efficiency and reduces residual feed and nutrient load. In S3, the integration of DO sensing, AI-empowered precision feeding, intelligent sludge removal, and ESG governance dashboards generates cumulative improvements in environmental, social, and governance performance. This finding is consistent with Garlock et al. (2024), who argue that aquaculture sustainability should be evaluated through integrated environmental, economic, and social indicators rather than production efficiency alone.

Table 2: Scenario Comparison of ESG Performance Impacts

Scenario	Environmental impact	Social impact	Governance impact
S0 Baseline	Higher water quality fluctuation and sludge risk	Higher labor intensity	Limited traceability
S1 DO Sensing	Improved oxygen stability	Reduced monitoring burden	Better environmental records
S2 DO + AI Feeding	Lower feed waste and nutrient load	Higher farmer trust	More transparent feeding decisions
S3 Full System	Cumulative ESG improvement	Reduced labor and risk	Auditable ESG indicators

5 Discussion

The findings suggest that AI-empowered smart aquaculture technologies should not be understood merely as automated devices, but as intelligent knowledge assets embedded with sensing data, algorithmic judgement, and expert knowledge. The three technologies of long-term dissolved oxygen sensing, AI-empowered precision feeding, and intelligent sludge removal can form a dynamic cycle of "data-knowledge-action-performance" through data accumulation, feedback correction, and decision-making transparency, which echoes IFKAD's focus on knowledge assets and sustainable organizational capabilities. The three core technologies and verification parameters set in this study also support the analysis direction of this paper to transform technical performance into ESG dynamic performance.

In terms of theoretical implications, this study integrates AIoT, smart aquaculture, system dynamics, and ESG governance to make up for the limitations of previous studies that focus on single-technology accuracy or short-term operational efficiency. AIoT literature points out that smart aquaculture can improve efficiency and sustainability through real-time environmental monitoring, data-driven decision-making, and automated control. System dynamics emphasizes feedback, delay, and nonlinear relationships, which is suitable for explaining the interaction between water quality, feeding, sediment, and fish growth (Huang & Khabusi, 2025; Sterman, 2000). Therefore, the contribution of this study is to illustrate how AI-empowered smart aquaculture technologies can generate cumulative ESG benefits through system structure, rather than only a single point of efficiency improvement.

In terms of management practices, farmers should avoid treating smart devices as siloed procurement items and instead establish integrated data governance and operational decision-making processes. Long-term sensing can reduce water quality uncertainty, AI feeding can improve feed utilization and reduce residual bait, and intelligent sludge removal can suppress sediment accumulation and delayed environmental risks. This integrated management logic is consistent with the Aquaculture Sustainability Assessment study, which suggests that aquaculture performance needs to consider environmental, economic, and social outcomes rather than just yield (Garlock et al., 2024).

In terms of ESG governance, this study suggests that the governance value of AI-empowered smart aquaculture comes from data traceability, auditable decisions, and transparent disclosure of key performance indicators (KPIs). Corporate sustainability research shows that highly sustainable organizations typically have well-established governance processes, stakeholder communication, and long-term performance orientation (Eccles et al., 2014). Therefore, the ESG outcomes of smart aquaculture do not occur naturally but need to be institutionalized through standardized verification, data formats, audit mechanisms, and policy incentives.

6 Conclusion

6.1 Summary of findings

This study developed a system dynamics framework to explain how AI-empowered smart aquaculture technologies influence environmental, social, and governance (ESG) performance through feedback structures,

stock-and-flow relationships, and time-delay effects. The findings suggest that long-term dissolved oxygen (DO) sensing, AI-empowered precision feeding, and intelligent sludge removal do not create value merely as isolated technologies. Rather, they function as an integrated decision-support system that connects sensing data, algorithmic inference, operational control, and ESG performance. From a system dynamics perspective, this study shows that water quality stability, feeding efficiency, sludge accumulation, fish biomass, labor intensity, and data traceability are dynamically interrelated. This finding is consistent with prior studies emphasizing that aquaculture systems involve complex interactions among environmental variables, biological growth, water quality management, and operational interventions (Muralidhar et al., 2023; Udayakumar et al., 2025).

6.2 Theoretical contribution

Theoretically, this study contributes to the literature by integrating AIoT, smart aquaculture, system dynamics, and ESG governance into a unified analytical framework. Previous studies have shown that AIoT can support real-time environmental monitoring, data-driven decision-making, and automated aquaculture management (Huang & Khabusi, 2025), while system dynamics is particularly suitable for analyzing nonlinear relationships, feedback loops, and delayed system behavior (Forrester, 1961; Sterman, 2000). This study extends these perspectives by conceptualizing AI-empowered smart aquaculture technologies as knowledge-intensive assets. These technologies transform sensing data, algorithmic inference, and expert rules into operational decisions and auditable ESG evidence. In this sense, the study also responds to IFKAD's knowledge asset perspective by showing how intelligent technologies can support sustainable organizational capabilities and long-term value creation.

6.3 Managerial and policy implications

From a managerial and policy perspective, the findings suggest that aquaculture operators should not treat smart devices as isolated equipment investments. Instead, DO sensing, AI-empowered precision feeding, intelligent sludge removal, and ESG governance dashboards should be implemented as an integrated management architecture. Such integration can reduce water quality uncertainty, improve feed utilization, lower sludge-related environmental risks, reduce labor burden, and strengthen data traceability. This implication is consistent with aquaculture sustainability research, which argues that aquaculture performance should be evaluated through integrated environmental, economic, and social indicators rather than production output alone (Garlock et al., 2024). For policymakers, the study suggests the need to support standardized verification protocols, interoperable data formats, ESG-oriented performance indicators, and incentive mechanisms that guide farmers from technology adoption toward long-term sustainability governance. This also aligns with corporate sustainability research, which indicates that sustainable performance depends on governance processes, stakeholder communication, and long-term organizational orientation (Eccles et al., 2014).

6.4 Limitations and future research

This study has several limitations. First, the model is primarily based on conceptual system dynamics modelling, selected field-verified technical parameters, literature-based assumptions, and expert judgment. Although the model structure was examined through structural verification, dimensional consistency checks, extreme-condition tests, and sensitivity analysis, future research should incorporate longitudinal field data and cross-site empirical validation to strengthen model robustness (Barlas, 1996). Second, the ESG performance indicator in this study remains a conceptual composite measure; future studies may refine the weighting and measurement of environmental, social, and governance dimensions through expert elicitation, field calibration, or multi-criteria decision analysis. Third, future research may combine system dynamics with life-cycle assessment, cost-benefit analysis, or agent-based modelling to further evaluate the economic feasibility, environmental impacts, farmer adoption behavior, and governance value of AI-empowered smart aquaculture systems across different species, farming scales, and production environments.

7 Acknowledgements

This study forms part of the research outcomes of the Ministry of Agriculture science and technology project "Innovative Aquaculture Long-Term Sensing Chip and Verification of Intelligent Pollutant Discharge" under project number 114 Agricultural Sciences-16.1.1-Water-01. The authors gratefully acknowledge the Ministry of Agriculture for its research funding and guidance.

8 Ethics Statement

This study did not involve human subjects, personal data, or animal experimentation conducted directly by the authors. The analysis was based on aggregated technical validation parameters, literature review, and system dynamics modelling. Therefore, formal human-subject ethical approval was not required.

9 AI Declaration

ChatGPT was used to assist with language refinement, literature identification, and figure drafting. All cited references were checked by the authors for relevance to the arguments of this study. The authors remain fully responsible for the research design, analysis, interpretation, and final content of the paper.

References

Al-Abri, S., Keshvari, S., Al-Rashdi, K., Al-Hmouz, R., & Bourdoucen, H. (2025). Computer vision-based approaches for fish monitoring systems: A comprehensive study. *Artificial Intelligence Review*, *58(6)*, *Article 185.*

Barlas, Y. (1996). Formal aspects of model validity and validation in system dynamics. *System Dynamics Review*, *12(3)*, 183-210.

Eccles, R. G., Ioannou, I., & Serafeim, G. (2014). The impact of corporate sustainability on organizational processes and performance. *Management Science*, *60(11)*, 2835-2857.

Forrester, J. W. (1961). Industrial dynamics. MIT Press.

Garlock, T. M., Asche, F., Anderson, J. L., Eggert, H., Anderson, T. M., Che, B., ... & Tveteras, R. (2024). Environmental, economic, and social sustainability in aquaculture: the aquaculture performance indicators. *Nature Communications*, *15(1)*, 5274.

Huang, Y. P., & Khabusi, S. P. (2025). Artificial intelligence of things (AIoT) advances in aquaculture: a review. *Processes*, *13(1)*, 73.

Muralidhar, M., Kumar, J. A., Suvana, S., Jayanthi, M., Vishwajeet, P., & Dayal, J. S. (2023). Dynamic modelling of coastal aquaculture systems: a review. *Aquatic Ecosystem Health & Management*, *26(3)*, 40-52.

Sterman, J. D. (2000). Business dynamics: Systems thinking and modeling for a complex world. McGraw-Hill.

Tacon, A. G., & Metian, M. (2008). Global overview on the use of fish meal and fish oil in industrially compounded aquafeeds: Trends and future prospects. *Aquaculture*, *285(1-4)*, 146-158.

Udayakumar, R., Kadirov, I., Radjabova, D., Fallah, M. H., Tursunov, M., & Masalieva, O. (2025). A system dynamics model for water quality management in recirculating aquaculture systems (RAS). *Natural and Engineering Sciences*, *10(2)*, 434-446.

Wei, Y., Jiao, Y., An, D., Li, D., Li, W., & Wei, Q. (2019). Review of dissolved oxygen detection technology: From laboratory analysis to online intelligent detection. *Sensors*, 19(18), 3995.

Xiao, Y., Huang, L., Zhang, S., Bi, C., You, X., He, S., & Guan, J. (2025). Feeding behavior quantification and recognition for intelligent fish farming application: A review. *Applied Animal Behaviour Scienc*e, *285*, 106588.

A System Dynamics Analysis of AI-Driven Competency Development in Business Schools

Chiung-Chuan Lo
Department of Business Administration, Shih Chien University. Taipei, Taiwan

Abstract

As artificial intelligence (AI) becomes increasingly embedded in business decision-making, analytics, marketing, finance, human resource management, and supply chain operations, business schools face growing pressure to redesign talent development models. AI adoption in business education should not be understood merely as the introduction of digital tools; rather, it represents a systemic transformation of curriculum design, learning analytics, feedback mechanisms, and competency assessment. This study develops a System Dynamics (SD) model to examine how AI-integrated learning environments influence competency development in business schools through the Knowledge, Skills, and Abilities (KSA) framework. In this model, Knowledge refers to students' accumulation of AI concepts, data analytics, management theories, and business decision-making knowledge; Skills refer to students' ability to use AI tools for data processing, case diagnosis, decision simulation, and project execution; and Abilities refer to students' higher-order capacity to integrate knowledge and skills in complex business contexts, including problem-solving, strategic judgment, cross-domain application, and responsible AI-enabled decision-making.

Methodologically, this study applies SD to construct causal loop diagrams and stock-and-flow diagrams that capture the feedback mechanisms, time delays, and nonlinear relationships underlying AI-driven competency development. Four key feedback loops are identified: the R1 Learning Accumulation Loop, which explains how AI feedback quality enhances learning engagement and knowledge acquisition; the R2 Skill Practice Loop, which illustrates how AI simulation tools and repeated practice strengthen skill development; the R3 Ability Reinforcement Loop, which shows how reflective feedback, peer interaction, project implementation, and teacher guidance deepen applied abilities; and the B1 Cognitive Load Loop, which explains how excessive course complexity and AI tool burden may increase cognitive load and learning fatigue.

This balancing mechanism is consistent with cognitive load theory, which argues that excessive task complexity consumes limited working memory resources and constrains learning effectiveness (Sweller, 1988).

Scenario-based model analysis compares three instructional designs: Traditional Teaching, AI-Assisted Learning, and AI-Integrated Experiential Learning. The findings suggest that traditional teaching supports knowledge accumulation but provides limited skill and ability development. AI-assisted learning improves knowledge acquisition and skill practice, yet may remain tool-oriented without project-based reflection. AI-integrated experiential learning produces the most balanced and sustained development across KSA by combining AI-assisted analysis, case simulation, enterprise projects, peer interaction, reflective journals, and teacher feedback. The study contributes to AI education, competency-based learning, and knowledge management (KM) by reframing business education as a dynamic, feedback-driven competency development system.

Keywords: Artificial Intelligence in Education; Business Schools; System Dynamics; KSA Framework; Competency Development.

Paper type: Academic Research Paper

1 Introduction

1.1 Background

As artificial intelligence (AI) rapidly enters the fields of corporate decision-making, business analysis, marketing management, financial technology, human resources, and supply chain management, business schools are under pressure to redefine talent cultivation models. AI not only transforms corporate knowledge creation and decision-making processes but also changes instructional design, learning analytics, personalized learning, and competency assessment methods in higher education. In recent years, AI in Education research has shown that AI has been widely used in adaptive learning, intelligent tutoring, learning prediction, intelligent assessment, and educational management, not only to improve learning efficiency but also to support learning process tracking and personalized feedback (Mahmud et al., 2025; Wang et al., 2024).

For business schools, AI-implemented courses are not just about adding digital tools but about the overall transformation of students' knowledge, skills, and abilities. AI empowerment has shifted learning from memorization and standardized tests to practical skills, knowledge application, and cognitive development

(Radu et al., 2024). Traditional courses often struggle to capture the dynamic feedback relationships among conceptual knowledge, practical skills, applied abilities, and learning performance. Therefore, this study adopts the Knowledge, Skills, and Abilities (KSA) framework to explore how AI-integrated learning environments affect student competency development in business schools (Caratozzolo et al., 2023; Faham et al., 2017).

1.2 Awareness of the problem

Existing AI education research mostly focuses on learning outcomes, learning satisfaction, personalized learning, or platform functions (Wang et al., 2024), but rarely explains from a systematic perspective how AI affects students' ability formation in the long term through feedback, delay, accumulation, and non-linear relationships. Competency-oriented education emphasizes that learning outcomes should not be limited to knowledge retention, but should be demonstrated as transferable, applicable, and measurable professional abilities (Frank et al., 2010; Voorhees, 2001). However, in AI-driven business education scenarios, students may experience a gap of "rapid knowledge acquisition" but "lack of ability deepening". While generative AI and learning analytics tools can enhance learning convenience, the lack of reflection, practice, and teacher guidance can also lead to cognitive load, tool dependence, or surface learning (Jia et al., 2025; Lee et al., 2025).

System Dynamics is suitable for dealing with such educational problems because it can analyze feedback loops, competency stock, delay effects, and balancing mechanisms. Faham et al. (2017) used system dynamics to analyze sustainability development in higher education, pointing out that student competency development involves multiple reinforcement and balancing loops, and can understand long-term change mechanisms through model simulations. Therefore, to understand the true impact of AI on business school competency development, it is necessary to move beyond single-point effectiveness evaluation and move toward dynamic system analysis.

1.3 Objectives of the study

Based on the above background and awareness of the problem, this study aims to construct a systematic dynamic model of "AI-driven business school competency development" and analyze how AI tools, learning engagement, intelligent feedback, project implementation, peer interaction, and reflective learning jointly affect KSA competency accumulation. Specifically, the objectives of this study include: first, to clarify the impact path of AI-enabled learning environment on the Knowledge, Skills, and Abilities of business school students; second, to establish causal loop diagrams and stock-and-flow diagrams to present key mechanisms such as learning accumulation, skill practice, ability reinforcement, and cognitive load; third, to compare the impact of different AI-assisted teaching strategies on students' long-term ability development, curriculum adaptability, and institutional innovation; and fourth, to put forward theoretical and practical suggestions for business schools to design AI-integrated curriculum and competency-based learning systems.

The expected contribution of this study is to integrate AI education, KSA competency framework, and System Dynamics to fill the gap in existing research that rarely deals with the dynamic mechanism of AI-driven competency development, and to reunderstand business school education as a capability development system with feedback, cumulative, and knowledge asset transformation characteristics.

2 Literature review

2.1 AI education and higher education transformation

AI in education has evolved from early computer-assisted teaching to integrated systems supporting adaptive learning, intelligent tutoring, learning analytics, prediction, assessment, and educational management. Recent studies (Wang et al., 2024) show that AI is no longer just a supporting tool but a core foundation reshaping instructional design and learning processes.

In higher education, AI enhances not only learning efficiency but also provides real-time feedback, personalized learning paths, and data-driven insights into student engagement and competency development. AI-powered intelligent assistants (Sajja et al., 2024) enable students to learn at their own pace while supporting knowledge construction and task-based learning. This shift transforms AI from a tool into a learning system, making education more dynamic and data-informed (Jia et al., 2025; Lee et al., 2025).

However, integrating AI also raises critical ethical and social concerns. Issues such as data privacy, algorithmic bias, academic integrity, and changing teacher–student roles must be addressed (Al-Zahrani & Alasmari, 2024). Therefore, AI in education should not only be evaluated by learning outcomes but also by its impact on learning interaction, competency development, and educational systems.

2.2 Business School Capability Development and AI-Driven Talent Cultivation

Business schools have traditionally focused on cultivating management knowledge, decision-making skills, and leadership capabilities. However, with the integration of AI into business operations, required competencies have expanded to include data analytics, AI tool application, digital transformation, cross-domain communication, ethical judgment as part of higher-order abilities, and complex problem-solving. Wu et al. (2026) highlight that AI reshapes business education across curriculum design, teaching methods, competency requirements, and governance, emphasizing the need for programs to align with evolving business functions.

From a competency-based perspective, learning outcomes should extend beyond knowledge acquisition to include analysis, application, and action. AI supports this shift by enabling simulated decision-making, data-driven case analysis, generative AI-assisted reporting, and real-time feedback. These tools provide students with more opportunities for practice and contextual learning, enhancing their ability to apply knowledge in real-world scenarios. However, AI integration also presents challenges, including over-reliance on tools, superficial learning, and increased cognitive load. Therefore, effective AI-enabled education must focus on how students build, transform, and internalize competencies through continuous action and reflection (Lee et al., 2025; Radu et al., 2024; Wang et al., 2024).

This study views AI-integrated learning environments as key drivers of competency development in business schools. It also argues that traditional course-based evaluation is insufficient, as it cannot capture the dynamic interaction between knowledge, skills, and performance. Instead, dynamic models are needed to better understand and assess long-term competency growth (Mahmud et al., 2025; Banihashem et al., 2022; Elsawah et al., 2022).

2.3 KSA Architecture: Knowledge, Skills, and Abilities

The KSA framework is widely used in education and talent development to describe the competencies required in professional contexts. It consists of three components: Knowledge, Skills, and Abilities. Knowledge refers to learners' understanding of concepts, theories, facts, and procedures. Skills represent the practical capability to perform tasks, use tools, analyze problems, and complete work activities. Abilities emphasize the integration of knowledge and skills in complex situations, reflecting higher-level competencies such as judgment, problem-solving, decision-making, and adaptability. Hlavac (2023) noted that KSA serves as an outcome-based framework linking education with observable performance.

In AI-driven business school education, these three components are highly interdependent. First, students must build a solid knowledge base in areas such as AI, data analytics, management theory, and digital transformation. Research suggests that AI education should be embedded within integrated competency frameworks rather than fragmented tool-based learning (Mahmud et al., 2025). Second, students need to develop practical skills, including applying AI tools for data processing, business analysis, decision simulation, and project execution. This aligns with industry demands for digital and analytical competencies (Caratozzolo et al., 2023).

Finally, Abilities are particularly critical, referring to students' capacity to apply knowledge and skills in uncertain and dynamic business environments. This includes strategic thinking, cross-domain integration, ethical judgment, and responsible decision-making. Sánchez-Ruiz et al. (2025) emphasize that AI competency must go beyond technical skills to include critical and ethical reasoning. Therefore, this study conceptualizes KSA as three core competency dimensions—Knowledge Base, Skill Set, and Abilities—essential for developing comprehensive managerial capabilities in the AI era.

2.4 System Dynamics and Educational Capacity Development

System Dynamics is well-suited for analyzing complex systems characterized by feedback loops, time delays, stock accumulation, and nonlinear changes. In education, competency development is not immediate but evolves over time through teaching input, learning interactions, practice, feedback, and reflection. Faham et al. (2017) applied SD to higher education for sustainable development and developed a model with multiple reinforcing and balancing loops to explain how students' competencies evolve over time. Similarly, Elsawah et al. (2022) emphasized that systems thinking in higher education helps students understand complex interrelationships, feedback structures, and dynamic problem situations, supporting the use of system-based approaches for analyzing educational competency development.

In this study, SD offers three key advantages. First, causal loop diagram (CLD) illustrate the relationships among AI tools, learning engagement, feedback quality, skill development, and ability formation (Sterman, 2000; Hamidi & Salavati, 2025). Second, stock-and-flow diagram (SFD) treat Knowledge Base, Skill Set, and Abilities as

dynamic competency stocks, enabling analysis of how competencies accumulate, interact, and change over time (Lin et al., 2020). Third, scenario-based model analysis allows comparison of different teaching approaches, including traditional instruction, AI-assisted learning, and AI-integrated experiential learning, and their long-term influence on competency development. In this sense, SD provides a useful methodological bridge between AI-enabled learning design and the dynamic analysis of Knowledge, Skills, and Abilities (KSA) development in business schools.

2.5 Summary: Research gaps

Existing research highlights AI's role in personalized learning, assessment, and analytics, while business education studies show that AI is reshaping required competencies. The KSA framework classifies competencies, and SD explains feedback and time-based changes. However, few studies integrate these perspectives to examine long-term competency development.

This study addresses this gap by combining KSA with SD to analyze how AI-enabled learning environments influence business students' competency growth. It focuses on dynamic processes such as learning accumulation, skill practice, ability reinforcement, and cognitive load, providing a more comprehensive understanding beyond short-term learning outcomes.

3 Research methodology

This study uses SD as the main research method, aiming to construct a dynamic model of AI-driven business school competency development. System Dynamics (SD) is suitable for analyzing complex systems with feedback loops, time delays, stock accumulation, and nonlinear changes. In educational research, SD has been applied to curriculum design, educational policy, and learning system analysis, helping researchers understand the hidden dynamic relationship between educational input and learning outcomes. Groff (2013) pointed out that education systems are highly interdependent and complex in policy, and SD can be used to analyze structural issues and future scenarios in education systems. Faham et al. (2017) also used SD to analyze sustainability development in higher education, showing that capacity development can be modeled and simulated through reinforcement and balance loops (Murthy et al., 2010).

3.1 Study design

This study adopts a research design that combines conceptual model construction and qualitative scenario-based comparison. First, based on the literature related to AI education, competency-based learning, and the KSA framework, the core variables affecting business school competency development are extracted. AI education research indicates that personalized and adaptive learning can enhance students' learning engagement, academic performance, and learning experience, especially in higher education (Du Plooy et al., 2024). Second, this study distinguishes students' competencies into three dimensions: Knowledge, Skills, and Abilities, and analyzes how AI tools, learning analytics, intelligent feedback, project implementation, peer interaction, and reflective learning affect competency accumulation. This study focuses on KSA and analyzes the expected long-term impact of AI-integrated curriculum through CLD, SFD, and qualitative scenario-based comparison.

3.2 Causal circuit diagram design

This study applies a causal loop diagram (CLD) to illustrate feedback relationships in AI-driven competency development within business schools. CLD, a key tool in System Dynamics, helps identify reinforcing and balancing loops, making it suitable for analyzing interactions among learning input, feedback quality, ability growth, and limiting factors. Prior studies (Groff, 2013; Faham et al., 2017) confirm that CLD effectively reveals hidden causal structures and dynamic mechanisms in educational systems.

The model includes four main loops. The R1 Learning Accumulation Loop shows how improved AI feedback enhances engagement and knowledge growth. The R2 Skill Practice Loop highlights how AI simulations and repeated practice strengthen skill development. The R3 Ability Reinforcement Loop explains how projects, reflection, and guidance support the integration of knowledge and skills into abilities. Finally, the B1 Cognitive Load Loop illustrates how excessive complexity and tool use may increase cognitive load, reducing engagement. This framework explains how AI influences long-term competency development through dynamic feedback processes.

3.3 Stock-and-Flow Diagram Design

In this study, the Stock-and-Flow Diagram (SFD) is used to further describe the accumulation and transformation of KSA competencies. Based on the Knowledge, Skills, and Abilities framework, the model sets three core stocks: Knowledge Base, Skill Set, and Abilities. Among them, the Knowledge Base is influenced by knowledge acquisition rate and is affected by knowledge forgetting or knowledge obsolescence. Skill Set accumulates from the skill development rate and may decline due to skill obsolescence. Abilities are formed by the ability deepening rate, which represents students' problem-solving, strategic judgment, and situational application abilities after integrating knowledge and skills. SFD can present the relationship between competency stock, flow, and time delay, which is helpful for analyzing long-term competency development (Sterman, 2000). Faham et al. (2017) also confirmed that SFD can be used to represent and, when parameterized, simulate long-term dynamics and feedback structures in higher education competency development.

3.4 Scenario design

This study compares three teaching scenarios (Table 1). Scenario 1: Traditional Teaching focuses on lectures, textbook learning, and case discussions, mainly supporting steady knowledge accumulation. Scenario 2: AI-Assisted Learning incorporates AI-based assessments, learning analytics, personalized materials, and real-time feedback to enhance learning efficiency and skill development. Scenario 3: AI-Integrated Experiential Learning combines AI tools with case simulations, corporate projects, peer interaction, reflective learning, and instructor guidance, aiming to promote comprehensive competency growth.

Table 1: Scenario Design

Scenario	Instructional Design Features	AI Level *	Expected KSA Development Logic
Scenario 1: Traditional Teaching	Teacher lectures, textbook reading, general case discussions, and final assessments	Low	Supports basic Knowledge accumulation, but provides limited opportunities for Skills practice and Abilities development
Scenario 2: AI-Assisted Learning	AI-assisted quizzes, learning analytics, personalized materials, intelligent feedback, and automated suggestions	Medium	Enhances Knowledge acquisition and Skills practice, but Abilities development may remain limited without projects and reflection
Scenario 3: AI-Integrated Experiential Learning	AI-assisted analysis, case simulations, enterprise projects, peer interaction, reflective journals, and teacher feedback	High	Promotes balanced KSA development through action, reflection, feedback, and competency transformation

* AI Level : AI Integration Level

4 Model Analysis and Scenario Comparison

This study constructs an AI-driven competency development model for business schools based on System Dynamics, and the model analysis can be divided into three part: first, proposing a competency development model with KSA as the core; second, identifying the main feedback loops in the AI-assisted learning environment; and third, comparing the expected impact of different teaching designs on students' long-term ability growth through qualitative scenario-based model analysis. Overall, the model analysis suggests that the impact of AI on business school students' ability development is not a single path or immediate effect, but is gradually formed through multiple mechanisms such as knowledge accumulation, skill implementation, reflective feedback, and ability reinforcement.

4.1 AI-driven business school capability development model

This study proposes an AI-driven business school competency development model based on the KSA framework (Knowledge, Skills, Abilities), illustrating how AI-integrated learning environments support the accumulation and transformation of student competencies. AI has been widely applied in adaptive learning, intelligent assessment, and learning analytics, enhancing personalized learning, feedback quality, and learning diagnosis (Wang et al., 2024). Furthermore, AI is reshaping business education by transforming curriculum design, competency requirements, and talent development models (Wu et al., 2026).

In this model, the Knowledge Base refers to students' understanding of AI concepts, data analytics, management theory, and digital transformation. AI-supported materials, feedback systems, and analytics help improve learning efficiency by enabling students to identify gaps and adjust their learning more quickly. This

aligns with the view that AI education should be implemented through integrated competency frameworks rather than fragmented tools (Mahmud et al., 2025).

The Skill Set represents students' ability to apply AI tools in data processing, business analysis, decision simulation, and project execution. AI simulations and generative tools provide repeated practice opportunities, allowing students to transform theoretical knowledge into practical and analytical skills. This reflects the need to align education with digital workforce competencies (Caratozzolo et al., 2023).

Abilities represent higher-level competencies, including problem-solving, strategic thinking, cross-domain collaboration, and responsible AI application in complex business contexts. Knowledge forms the foundation, skills enable application, and abilities integrate both into effective professional performance.

Overall, the model emphasizes that AI does not directly create abilities but influences competency development through three mechanisms: enhancing knowledge accumulation via feedback, strengthening skills through practice and simulation, and fostering abilities through project implementation, reflection, and guidance. Thus, AI-enabled business education should be viewed as a dynamic, long-term system of competency development rather than a tool-based or course-specific intervention.

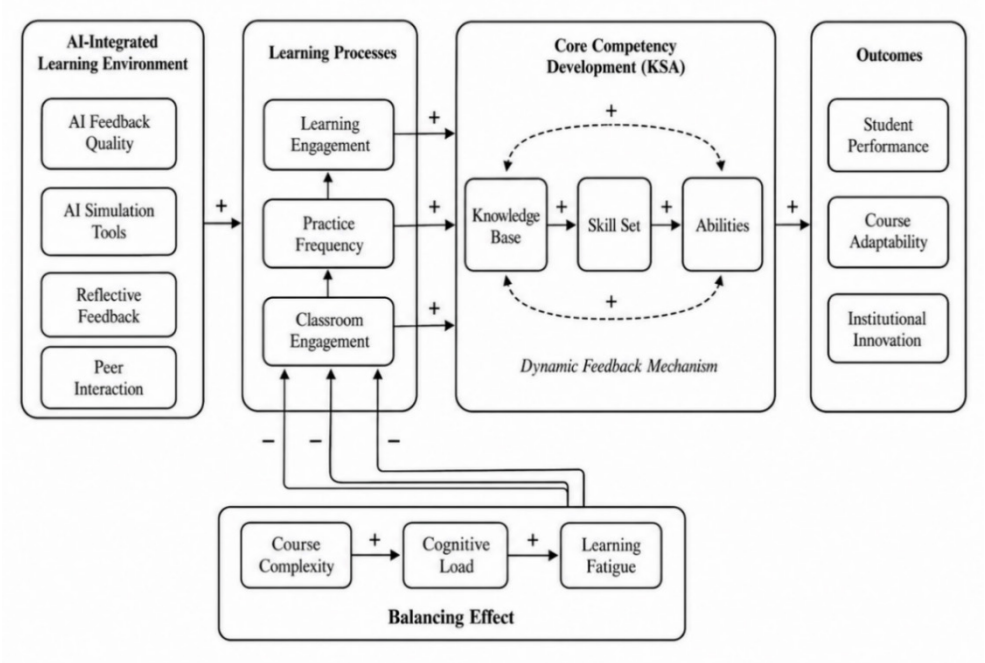


Figure 1: AI-driven business school capability development model diagram

4.2 Main feedback loop analysis

This study identifies three reinforcing loops and one balancing loop (Figure 2) to explain AI-driven competency development in business schools. The first is the R1 Learning Accumulation Loop, which highlights the positive relationship between AI feedback quality, learning engagement, and knowledge growth. AI tools—such as real-time quizzes, personalized recommendations, and learning analytics—help students identify knowledge gaps and adjust strategies, increasing engagement and accelerating knowledge acquisition. This process strengthens self-efficacy, further reinforcing continued learning and knowledge accumulation (Wang et al., 2024).

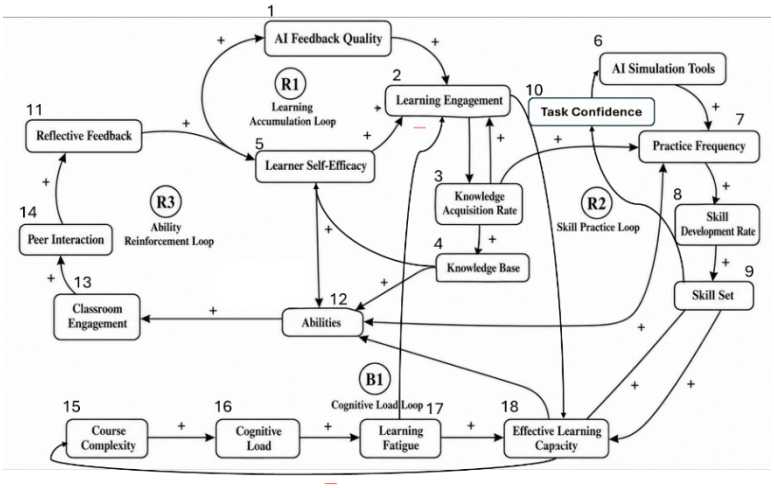


Figure 2: three reinforcing loops and one balancing loop

The second is the R2 Skill Practice Loop, which explains how AI simulation tools, repeated practice, and skill development reinforce each other. Through business simulations, data dashboards, and generative AI-assisted decision-making, students gain opportunities to apply knowledge in practical contexts. Repeated task execution enhances operational and analytical skills, aligning with research that emphasizes the role of project-based learning (PjBL) in improving competency through practice, collaboration, and reflection (Guo et al., 2020; Wu et al., 2026). The third is the R3 Ability Reinforcement Loop, focusing on the development of higher-level abilities. Through project implementation, reflective feedback, peer interaction, and teacher guidance, students integrate knowledge and skills into complex capabilities such as problem-solving, strategic thinking, and responsible AI application. This loop emphasizes that abilities are formed through continuous action and reflection rather than passive learning.

This mechanism is consistent with experiential learning and feedback research, which emphasizes that higher-order abilities emerge through experience, reflection, and feedback-supported improvement (Banihashem et al., 2022; Wijnen-Meijer et al., 2022).

Finally, the B1 Cognitive Load Loop represents a balancing mechanism. Excessive complexity in AI tools or course design may increase cognitive load and learning fatigue, reducing engagement and limiting skill and ability development. This highlights the importance of proper instructional design, including scaffolding, task segmentation, and guided support, to ensure sustainable long-term competency growth.

4.3 Scenario comparison outcomes

Based on the proposed system dynamics model, this study conducts a qualitative scenario-based model analysis (Table 2) to compare how different levels of AI integration influence students' Knowledge, Skills, and Abilities (KSA). Traditional Teaching mainly supports Knowledge accumulation through lectures, textbook reading, and case discussions, but provides limited opportunities for AI tool practice, project implementation, and ability transformation. AI-Assisted Learning improves Knowledge acquisition and Skills development through AI-assisted quizzes, learning analytics, personalized materials, and intelligent feedback; however, if learning remains tool-oriented, students' higher-order Abilities may remain limited.

Table 2: Scenario Comparison Outcomes

Scenario	Knowledge	Skills	Abilities	Overall Outcome
Traditional Teaching	Medium	Low to medium	Low	Helps students build basic knowledge, but the transformation from knowledge to skills and higher-order abilities remains insufficient
AI-Assisted Learning	High	Medium to high	Medium	Improves knowledge learning and skill development, but abilities may remain at the initial application level if learning is tool-oriented
AI-Integrated Experiential Learning	High	High	High	Produces the most balanced and sustained development across the three KSA dimensions

AI-Integrated Experiential Learning provides the strongest support for KSA development because it combines AI-assisted analysis, case simulation, enterprise projects, peer interaction, reflective journals, and teacher feedback. This scenario directly supports the R3 Ability Reinforcement Loop, in which project implementation, reflective feedback, peer collaboration, and teacher guidance help students integrate knowledge and skills into problem-solving, strategic judgment, cross-domain application, and responsible AI-enabled decision-making. Project-based learning improves higher education outcomes through practical tasks, collaboration, and reflection (Guo et al., 2020). Learning analytics strengthens feedback practices by helping students and teachers identify learning progress and competency gaps (Banihashem et al., 2022). Experiential learning also emphasizes that ability development occurs through experience, reflection, conceptualization, and active experimentation (Wijnen-Meijer et al., 2022). Therefore, AI-Integrated Experiential Learning is expected to produce the most balanced and sustainable development across the three KSA dimensions.

5 Discussion

5.1 Theoretical contributions

The main theoretical contribution of this study is the integration of AI education, the KSA competency framework, and System Dynamics to propose a dynamic model of competency development in business schools. While existing AI education research largely focuses on applications such as adaptive learning, intelligent assessment, and personalized learning (Wang et al., 2024), it rarely explains how AI influences long-term competency formation through feedback, time delays, and cumulative learning processes.

This study redefines KSA as Knowledge, Skills, and Abilities, conceptualizing them as three interdependent and evolving competency stocks rather than static categories. Drawing on Caratozzolo et al. (2023), it emphasizes that effective talent development in the digital era requires understanding how these components interact. In this model, the Knowledge Base supports understanding of AI, data analytics, and management theory; the Skill Set reflects the ability to apply AI tools in analysis and decision-making; and Abilities represent higher-order competencies such as problem-solving, strategic thinking, cross-domain integration, and responsible decision-making.

Furthermore, this study applies System Dynamics to address limitations of traditional competency research, which often overlooks feedback and temporal effects. Building on Faham et al. (2017), the model incorporates reinforcement loops (learning accumulation, skill practice, ability development) and a balancing loop (cognitive load). These illustrate that competency growth is nonlinear and influenced by factors such as feedback quality, practice frequency, reflection, and learning complexity.

Overall, this research advances AI education from a tool-oriented perspective to a dynamic framework that explains long-term competency development and knowledge accumulation.

5.2 Impact on knowledge management

From a KM perspective, this study views AI-driven business school competency development as the accumulation and transformation of educational knowledge assets. Within the KSA framework, Knowledge refers to understanding of AI, data analytics, and management concepts; Skills represent the ability to apply AI tools in analysis and decision-making; and Abilities reflect the integration of knowledge and skills into higher-level competencies such as problem-solving and strategic judgment.

AI implementation thus goes beyond tool adoption, functioning as a mechanism that converts learning data, project outputs, reflection records, and assessment results into reusable knowledge assets. Learning analytics further supports this process by collecting and visualizing data to identify learning progress, competency gaps, and improvement needs (Banihashem et al., 2022).

This knowledge management perspective aligns with IFKAD's concern with knowledge assets because students' learning traces, project outputs, competency records, and individual learning outcomes can be transformed into institutional knowledge assets that support curriculum redesign, teaching improvement, and long-term educational innovation.

5.3 Model interpretation

This research model points out that the impact of AI on business school competency development does not simply come from tool introduction, but gradually promotes the accumulation of Knowledge, Skills, and Abilities through feedback, implementation, reflection, and ability transformation mechanisms. AI feedback and learning analytics help form a knowledge base, while AI simulation tools, case studies, and project tasks can increase

practice frequency and promote skill set development. Furthermore, through project implementation, teacher guidance, and reflection activities, students can integrate knowledge and skills into abilities, demonstrating problem-solving, strategic judgment, and responsible decision-making abilities. However, if the course complexity and AI tool load are too high, it will form a cognitive load loop, reducing learning investment and inhibiting ability growth.

6 Conclusion

This study develops a systematic, dynamic model of AI-driven competency development in business schools based on the KSA framework. It argues that AI integration should not be viewed merely as the addition of digital tools, but as a feedback-driven system that enhances how competencies are accumulated, transformed, and deepened over time. AI technologies—such as adaptive learning, intelligent assessment, learner analytics, and prediction—support personalized learning and strengthen knowledge construction processes (Wang et al., 2024). The study's key theoretical contribution lies in integrating AI education, the KSA framework, and System Dynamics. Within this framework, Knowledge refers to students' understanding of AI, data analytics, management theory, and decision-making concepts. Skills represent the ability to apply AI tools for data processing, case analysis, and decision simulation. Abilities refer to higher-order competencies, including problem-solving, strategic thinking, cross-domain integration, and responsible decision-making in complex business contexts. This aligns with Caratozzolo et al. (2023), who emphasize the interconnected and dynamic nature of KSA in responding to digital transformation.

Methodologically, the study applies System Dynamics through four loops: the R1 Learning Accumulation Loop, R2 Skill Practice Loop, R3 Ability Reinforcement Loop, and B1 Cognitive Load Loop. These illustrate how AI feedback, engagement, practice, project-based learning, and cognitive load interact to influence long-term competency growth. As noted by Faham et al. (2017), System Dynamics effectively captures feedback mechanisms and time-based changes often overlooked in traditional education research. Practically, findings suggest that traditional teaching supports knowledge accumulation but has limited impact on skill and ability development. AI-assisted learning improves efficiency in knowledge acquisition and skill practice, yet without project-based tasks and reflective feedback, ability growth remains constrained. In contrast, AI-integrated experiential learning—combining simulations, real-world projects, peer interaction, and instructor guidance—promotes balanced development across KSA dimensions (Guo et al., 2020).

This study is limited by its conceptual nature and lack of large-scale empirical validation. Future research should incorporate longitudinal data, quantify model parameters, and compare different teaching scenarios. Overall, it emphasizes that AI education should focus on building sustainable, adaptive, and feedback-driven competency development systems rather than simply adopting new technologies.

7 Declarations

This study is a conceptual modeling study based on published literature and does not involve human participants, personal data, or experimental intervention. Google NotebookLM was used only to assist with literature organization and summarization. All arguments, interpretations, and final writing were reviewed and revised by the author.

References

Al-Zahrani, A. M., & Alasmari, T. M. (2024). Exploring the impact of artificial intelligence on higher education: The dynamics of ethical, social, and educational implications. *Humanities and Social Sciences Communications*, *11*(1), 912. https://doi.org/10.1057/s41599-024-03432-4

Banihashem, S. K., Noroozi, O., Van Ginkel, S., Macfadyen, L. P., & Biemans, H. J. (2022). A systematic review of the role of learning analytics in enhancing feedback practices in higher education. *Educational Research Review*, *37*, 100489. https://doi.org/10.1016/j.edurev.2022.100489

Caratozzolo, P., Azofeifa, J. D., Mejía-Manzano, L. A., Rueda-Castro, V., Noguez, J., Magana, A. J., & Benes, B. (2023, October). A matrix taxonomy of knowledge, skills, and abilities (KSA) shaping 2030 labor market. In *2023 IEEE Frontiers in Education Conference (FIE)* (pp. 1-8). IEEE.

Du Plooy, E., Casteleijn, D., & Franzsen, D. (2024). Personalized adaptive learning in higher education: A scoping review of key characteristics and impact on academic performance and engagement. *Heliyon*, *10*(21).

Elsawah, S., Ho, A. T. L., & Ryan, M. J. (2022). Teaching systems thinking in higher education. *INFORMS Transactions on Education*, *22*(2), 66-102.

Faham, E., Rezvanfar, A., Mohammadi, S. H. M., & Nohooji, M. R. (2017). Using system dynamics to develop education for sustainable development in higher education with the emphasis on the sustainability competencies of students. Technological Forecasting and Social Change, 123, 307-326. https://doi.org/10.1016/j.techfore.2016.03.023

Frank, J. R., Snell, L. S., Cate, O. T., Holmboe, E. S., Carraccio, C., Swing, S. R., ... & Harris, K. A. (2010). Competency-based medical education: theory to practice. *Medical Teacher*, *32*(8), 638–645.

Groff, J. S. (2013). Dynamic systems modeling in educational system design & policy. *Journal of New Approaches in Educational Research*, *2*(2), 72–81.

Guo, P., Saab, N., Post, L. S., & Admiraal, W. (2020). A review of project-based learning in higher education: Student outcomes and measures. *International Journal of Educational Research*, *102*, 101586.

Hamidi, A., & Salavati, S. (2025). Modeling system dynamics in computational thinking integration using causal loop diagrams. *Thinking Skills and Creativity*, 102003. https://doi.org/10.1016/j.tsc.2025.102003

Hlavac, J. (2023). Knowledge, skills and abilities (KSAs) as a metric to re-conceptualise aptitude: A multi-stakeholder perspective. *The Interpreter and Translator Trainer*, *17*(1), 29–53.

Jia, W., Pan, L., & Neary, S. (2025). Effect of GenAI dependency on university students' academic achievement: The mediating role of self-efficacy and moderating role of perceived teacher caring.. *Behavioral Sciences*, *15*(10), 1348.

Lee, H. P., Sarkar, A., Tankelevitch, L., Drosos, I., Rintel, S., Banks, R., & Wilson, N. (2025, April). The impact of generative AI on critical thinking: Self-reported reductions in cognitive effort and confidence effects from a survey of knowledge workers. In Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (pp. 1–22).

Lin, G., Palopoli, M., & Dadwal, V. (2020). From causal loop diagrams to system dynamics models in a data-rich ecosystem. In Leveraging data science for global health (pp. 77-98). Cham: Springer International Publishing.

Mahmud, M. M., Monib, W. K., Qazi, A., Wong, S. F., Ramachandiran, C. R., & Azizan, S. N. (2025). Developing AI education competency framework: A systematic literature review. *Open Praxis*, *17*(4), 730–748.

Murthy, S., Gujrati, R., & Iyer, S. (2010, December). Using system dynamics to model and analyze a distance education program. In *Proceedings of the 4th ACM/IEEE International Conference on Information and Communication Technologies and Development (pp. 1–8).*

Radu, C., Ciocoiu, C. N., Veith, C., & Cătălin, R. (2024). Artificial intelligence and competency-based education: A bibliometric analysis. *Amfiteatru Economic*, *26*(65), 220–240.

Sajja, R., Sermet, Y., Cikmaz, M., Cwiertny, D., & Demir, I. (2024). Artificial intelligence-enabled intelligent assistant for personalized and adaptive learning in higher education. *Information*, *15*(10), 596.

Sánchez-Ruiz, L. M., Llobregat-Gómez, N., Vega-Fleitas, E., & Moll-López, S. (2025). AI competency assessment and ranking: A framework for higher education. *Applied Sciences*, *15*(22), 12248.

Sterman, J. D. (2000). *Business dynamics: Systems thinking and modeling for a complex world*. Irwin/McGraw-Hill.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*(2), 257–285.

Voorhees, R. A. (2001). Competency-based learning models: A necessary future. *New Directions for Institutional Research*, *2001*(110), 5–13.

Wang, S., Wang, F., Zhu, Z., Wang, J., Tran, T., & Du, Z. (2024). Artificial intelligence in education: A systematic literature review. *Expert Systems with Applications*, *252*, 124167. https://doi.org/10.1016/j.eswa.2024.124167

Wijnen-Meijer, M., Brandhuber, T., Schneider, A., & Berberat, P. O. (2022). Implementing Kolb's experiential learning cycle by linking real experience, case-based discussion and simulation. Journal of Medical Education and Curricular Development, 9, 23821205221091511.

Wu, Q., Chen, L., Chen, M., & Huang, Y. (2026). Exploring the impact of artificial intelligence on business talent development in higher education: A systematic literature review and research agenda. *The International Journal of Management Education*, *24*(1), 101287. https://doi.org/10.1016/j.ijme.2025.101287

Evolution of Knowledge Landscapes in AI-KM: A Longitudinal Analysis of Divergence between Term Diversity and Community Attention (2015-2024)

Hsuan-Hao Lo, Sirirat Sae Lim
National Yang Ming Chiao Tung University, Hsinchu, Taiwan

Abstract

This study adopts a longitudinal bibliometric approach combined with co-word network analysis to examine the development of the AI–KM field over the past decade. While research at the intersection of artificial intelligence (AI) and knowledge management (KM) has expanded rapidly in recent years, existing bibliometric studies have remained largely cross-sectional and descriptive, offering limited insight into how the field's knowledge structure has evolved during this period of rapid growth. Drawing on 2,857 AI–KM publications indexed in the Web of Science between 2015 and 2024, this study investigates the evolution of the field through three complementary lenses: publication growth trajectories, conceptual cluster structures, and cross-period changes in author keywords. These analyses collectively reveal the overall trajectory and structural reconfiguration of the AI–KM field, identifying the continuity of core themes, shifts in research focus, and the emergence of new topics. By adopting a longitudinal perspective, this study addresses a key limitation of prior research in capturing dynamic changes in knowledge structure and extends understanding of the AI–KM field beyond a static snapshot toward an evolutionary view. The findings provide an empirical foundation for future theory building and research agenda setting in this area.

Keywords: Artificial Intelligence, Knowledge Management, Bibliometric Analysis, Co-word Network, Science Mapping

Paper type: Academic Research Paper

1 Introduction

In recent years, the intersection of artificial intelligence (AI) and knowledge management (KM) has emerged as a rapidly growing research domain (Cantu-Ortiz, 2021). Prior bibliometric studies have begun to map the overall landscape of this interdisciplinary field. For example, Cantu-Ortiz (2021), drawing on publications from 2011 to 2019, examined research trends and future agendas in AI and KM, while Kushwaha and Maurya (2025) complemented narrative reviews by analyzing core sources, authors, institutions, and co-word relationships. Collectively, these studies suggest that the AI–KM field has maintained strong growth momentum and offer initial insights into its major research themes, leading sources, and knowledge foundations.

However, existing research remains largely cross-sectional, focusing primarily on descriptive dimensions such as publication volume, citation performance, keyword co-occurrence, thematic clustering, and author productivity. Relatively few studies have adopted a longitudinal perspective to examine how the knowledge structure of the field has evolved during its period of rapid expansion. As a result, the developmental dynamics of the field remain insufficiently understood.

This gap is particularly notable given the highly interdisciplinary nature of AI–KM, which integrates knowledge from management, information science, engineering, and organizational studies. The rapid growth of a field does not necessarily imply the stable integration of its knowledge structure. As research output increases, the variety of author keywords and conceptual labels often expands as well, creating a surface impression of thematic diversification. Yet the expansion of terminological variety does not necessarily mean that newly introduced concepts are effectively integrated into the central structure of the field. Instead, rapid growth can coincide with the consolidation of core themes, the weakening of local clusters, or an uneven pace between the expansion of concepts and the redistribution of collective attention. In short, the apparent vibrancy of the AI–KM field does not in itself indicate greater structural integration.

Against this backdrop, this study draws on the notions of term diversity and community attention as a framework for examining the evolution of the AI–KM knowledge landscape. In this study, term diversity refers to the breadth and temporal variation of author keywords used in the field, reflecting the range of conceptual vocabulary through which researchers position their work. Community attention, by contrast, refers to the relative prominence and concentration of specific keywords across the field, indicating the thematic areas that attract sustained collective focus. Although these two dimensions are related, they may evolve differently. New

terms may continue to emerge while research attention remains concentrated on a limited number of established themes. Conversely, some terms may remain prominent even as their conceptual meaning shifts over time. Tracing the relationship between these two dimensions helps reveal how the AI–KM knowledge landscape has been reconfigured during its period of rapid growth.

Building on this framework, this study combines bibliometric analysis with co-word network analysis to examine the evolution of the AI–KM knowledge landscape across multiple layers. Beyond publication growth, the analysis examines the conceptual distribution of author keywords, the structural characteristics of the co-word network, and cross-period changes in thematic composition. By integrating these layers, this study traces how conceptual linkages in the AI–KM field have formed, expanded, and reorganized over time. Accordingly, this study addresses the following questions:

1. *What stage-wise characteristics can be identified in the publication growth trajectory of the AI–KM field between 2015 and 2024?*
2. *What cluster patterns and core themes emerge in the conceptual structure formed by author keywords in this field?*
3. *How have the high-frequency author keywords of the field shifted across different developmental periods?*

This study makes three main contributions. First, it complements existing bibliometric research by adopting a longitudinal perspective, thereby extending understanding of the field beyond a static snapshot toward an evolutionary view. Second, based on publications from 2015 to 2024, it integrates descriptive analysis, co-word network analysis, and cross-period keyword comparison within a multi-layered bibliometric framework to reveal patterns in the field's structural development. Third, it provides an empirical foundation for future theory building and research agenda setting in the AI–KM domain.

2 Methodology

This study employs a quantitative research design that combines longitudinal bibliometric analysis with co-word network analysis and network structural indicators to examine the evolution of the AI–KM knowledge landscape between 2015 and 2024. These approaches have been widely used to identify research themes, map knowledge structures, and trace their long-term evolution across fields such as knowledge management, information systems, AI innovation, and digital innovation (Bernatović et al., 2022; Cobo et al., 2011; Gaviria-Marin et al., 2019). They are particularly suitable for a field such as AI–KM, characterized by rapid growth, interdisciplinary integration, and continually shifting knowledge structures.

Early literature reviews of emerging fields often relied on narrative synthesis, qualitative induction, or descriptive summaries to organize themes, theoretical foundations, and key findings (Grant & Booth, 2009). While useful for establishing an initial understanding of a field, these approaches become less effective as research output grows rapidly, making it difficult to capture structural linkages among themes, internal reorganization processes, and evolutionary changes over time. To address these limitations, later studies increasingly adopted systematic literature reviews and bibliometric methods (Donthu et al., 2021; Tranfield et al., 2003), using citations, co-word relationships, and network structures to map research themes and knowledge development over time (Cobo et al., 2011; Donthu et al., 2021).

Building on this methodological trajectory, the present study combines longitudinal bibliometric analysis and co-word network analysis to examine the developmental profile, conceptual linkages, and structural shifts of the AI–KM field.

2.1 Research Design

This study employs bibliometric analysis as its primary research method to systematically map the developmental profile, core concepts, and thematic evolution of the AI–KM field between 2015 and 2024. Compared with traditional narrative reviews, bibliometric analysis offers a more systematic, objective, and visual approach to examining the knowledge structure, conceptual distribution, and developmental patterns of a research field. To enhance the transparency and reproducibility of the research process, this study follows the methodological guidance of Donthu et al. (2021) and Öztürk et al. (2024), proceeding through data retrieval, data cleaning, network construction, analysis, and visualization.

Analytically, the study focuses on conceptual structure analysis within a science mapping framework (Hoang, 2025), using keyword co-occurrence analysis and cross-period thematic comparison as its main techniques. The former is used to identify co-occurrence relationships and thematic clusters among core concepts, while the latter is used to trace the continuity, transformation, emergence, and decline of themes across different periods.

By adopting a longitudinal perspective, this design helps reveal how knowledge in the AI–KM field has accumulated, diversified, and reorganized over time.

2.2 Data Collection

This study uses the Web of Science Core Collection (WoS) as its primary data source. WoS is well suited to longitudinal bibliometric and network analyses because of its broad interdisciplinary coverage, rigorous editorial selection and quality standards, and comprehensive citation and bibliographic records. Accordingly, it has been widely adopted as a primary data source in bibliometric research (Gaviria-Marin et al., 2019; Merigó et al., 2015; Wamba et al., 2021).

The search was conducted on January 30, 2026, and covered the period from January 1, 2015 to December 31, 2024. This period was selected because the AI–KM field has entered a phase of marked growth over the past decade, and the 2015–2024 window provides a relatively complete interval for observing the formation and expansion of its knowledge landscape. To reduce potential bias arising from the limited citation accumulation of very recent publications, records from 2025 were excluded from the core analysis.

The search query was constructed by combining AI-related and KM-related terms with the Boolean operators "OR" and "AND" and was applied to the Title (TI), Abstract (AB), and Author Keywords (AK) fields (Donthu et al., 2021; Hoang, 2025; Marzi et al., 2025). AI-related terms included *artificial intelligence*, *machine learning*, *deep learning*, and *generative AI*, while KM-related terms included *knowledge management*, *knowledge sharing*, *knowledge transfer*, and *knowledge creation*.

Document types were restricted to journal articles and conference papers, and the language was limited to English to ensure consistency, comparability, and overall research quality. The initial search returned 3,521 records. After excluding records that lacked essential information, such as publication year or author keywords, as well as non-English entries, the final dataset consisted of 2,857 publications.

2.3 Data Cleaning and Author Keyword Normalization

To improve the accuracy of the keyword co-occurrence analysis, this study conducted a systematic cleaning and normalization of the author keywords extracted from the WoS records. In the raw data, author keywords often exhibit inconsistencies in capitalization, variations between abbreviations and full forms, singular–plural differences, and synonymous expressions. Without preprocessing, the same concept may be split across multiple nodes, thereby distorting the structure of the co-word network and the formation of thematic clusters (Cobo et al., 2011; Muñoz-Leiva et al., 2012). To address these issues, a four-stage cleaning procedure was implemented, consisting of parenthetical removal, synonym consolidation, abbreviation expansion, and singular–plural unification, supplemented by a stopword list and a residual singular–plural scan to improve overall data quality.

The first stage removed parenthetical content. Some authors include supplementary notes or abbreviations in parentheses, such as "knowledge transfer (KT)." Although useful for readers, this practice creates multiple surface forms of the same concept across documents. Parenthetical content was therefore removed, standardizing such cases to a single form and reducing node fragmentation.

The second stage consolidated synonymous expressions. Terms with equivalent meanings but different spellings were merged into a unified form. For example, both "information and knowledge management" and "information & knowledge management" were standardized as "knowledge management." These merging rules were developed based on conventional usage in the domain literature and further verified through manual review to avoid over-merging concepts that differ in abstraction or scope.

The third stage expanded common abbreviations to prevent a single concept from being split across multiple nodes because of surface variation. For example, "AI" was expanded to "artificial intelligence," "LLM" to "large language model," and "CNN" to "convolutional neural network." Expansion rules were applied only when the abbreviation had a clear one-to-one mapping. In ambiguous cases, such as "RL," which may refer to either reinforcement learning or representation learning, expansion was determined case by case based on contextual cues in order to avoid erroneous merging.

The fourth stage unified singular and plural forms by mapping plurals onto their singular counterparts, such as "neural networks" to "neural network" and "large language models" to "large language model." This step is especially important in co-occurrence analysis, as singular–plural variation may underestimate the frequency of high-occurrence terms and obscure core themes.

In addition, this study compiled a stopword list to exclude overly generic terms that did not meaningfully contribute to conceptual structure analysis, such as "training," "optimization," and "feature extraction."

Although these terms appear frequently, their broad semantic scope may introduce large numbers of non-substantive linkages into the network and reduce the interpretability of thematic clusters.

Finally, to verify the quality of the cleaning process, a residual singular–plural scan was conducted on all keywords with a frequency of four or more. Variants not captured by the general rules were manually identified and added to the rule set, resulting in more than twenty additional normalization rules covering terms such as "transformers," "support vector machines," "expert systems," and "foundation models." Following this procedure, noise in the original keyword set was substantially reduced, improving the stability and interpretability of the subsequent co-occurrence and thematic structure analyses (Cobo et al., 2011; Hoang, 2025).

2.4 Data Analysis

Following the general procedures of science mapping and bibliometric analysis, the data workflow of this study comprised four stages: data retrieval, preprocessing, network extraction, and analysis and visualization. Bibliographic records meeting the search criteria were first downloaded from the WoS Core Collection. The records were then cleaned, deduplicated, and filtered to retain the fields required for analysis, including publication year, source, author and institutional information, and author keywords. Records lacking any of these fields were excluded. The final dataset served as the analytical basis of this study.

The cleaned dataset was subsequently imported into VOSviewer for network visualization and analysis. VOSviewer is widely used in bibliometric research and is particularly well suited to keyword co-occurrence analysis. Through cluster analysis and network indicators such as links and total link strength, it supports the identification of core concepts, thematic clusters, and the broader knowledge structure of a research field (Bernatović et al., 2022; Van Eck & Waltman, 2010). Using author keywords as the unit of analysis, this study conducted co-occurrence analysis to examine the relational connectivity and structural prominence of keywords within the overall conceptual network.

To examine temporal shifts in research focus, this study further compared high-frequency keywords across three developmental periods based on the growth trajectory of the field. For each period, keyword frequency was normalized as the proportion of documents containing a given keyword relative to the total number of documents in that period, and cross-period differences were calculated in percentage points. This comparison provides a basis for tracing the continuity, weakening, and emergence of themes across different stages of the AI–KM field's development, and supports the subsequent interpretation of the results.

3 Results and Discussion

This section presents the empirical results for the AI–KM field between 2015 and 2024. Following the analytical logic of bibliometric analysis and science mapping, the discussion proceeds from a descriptive overview to the conceptual structure and cross-period thematic comparison of the field.

As shown in Figure 1, the bars represent annual publications, the solid line denotes the three-year moving average, and the dashed line indicates cumulative publications. Between 2015 and 2024, the annual output of the AI–KM field exhibited a clear upward trend, increasing from 72 publications in 2015 to 782 in 2024. The cumulative curve confirms the sustained expansion of the field, while the three-year moving average indicates a marked acceleration in growth after 2022, suggesting that AI–KM research has entered a phase of rapid expansion.

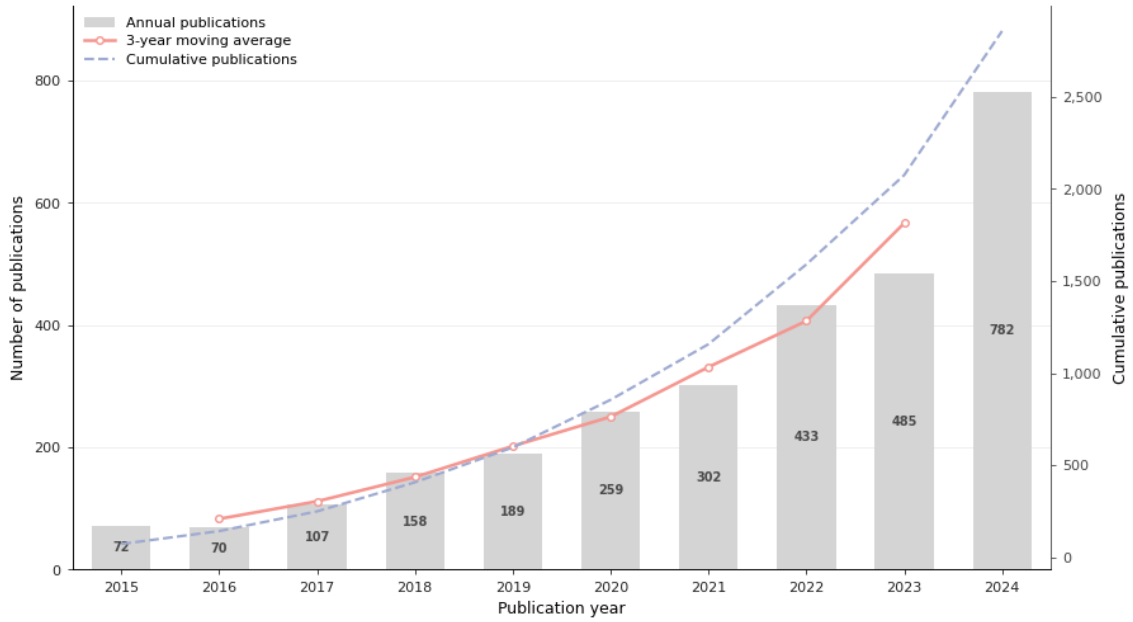


Figure 1: Annual and cumulative publication trends in AI–KM research (2015–2024

3.1 Conceptual Structure

To identify the conceptual structure of AI–KM research, this study employs co-word analysis based on author keywords. Co-word analysis captures the co-occurrence of terms within the same document and is widely used to depict conceptual linkages, thematic clusters, and knowledge structures in science mapping research (Cobo et al., 2011; Jiang et al., 2023). Although it provides a systematic way to reveal associative patterns among concepts, its results are highly sensitive to the quality, consistency, and normalization of author keywords and therefore require careful contextual interpretation (Muñoz-Leiva et al., 2012).

VOSviewer identified a total of 8,365 author keywords in the dataset. To improve the interpretability of the network and reduce noise from low-frequency terms, the minimum occurrence threshold was set at 12, resulting in 106 keywords being included in the analysis and yielding three major clusters. Based on occurrence, link count, and total link strength, the core nodes of the network include *knowledge transfer*, *deep learning*, *transfer learning*, *artificial intelligence*, *and knowledge management*, as shown in Figure 2 and summarized in Table 1.

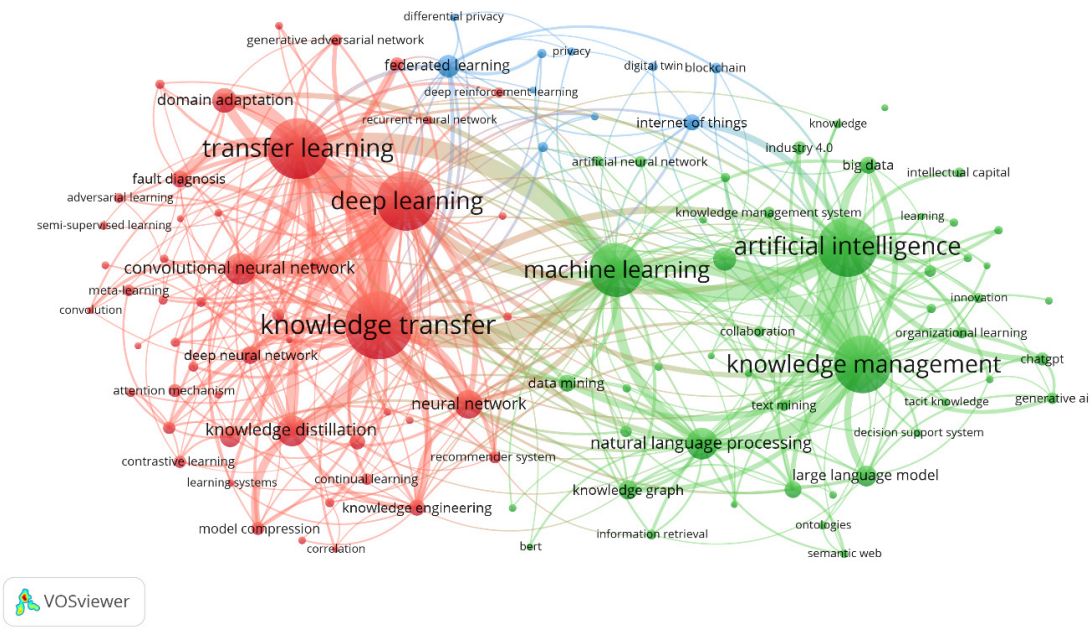


Figure 2: Conceptual structure of AI–KM research based on author keyword co-occurrence (2015–2024)

Cluster 1 centers on AI-driven knowledge transfer and learning mechanisms, with *knowledge transfer*, *transfer learning*, *deep learning*, *and knowledge distillation* as its core nodes. This cluster reflects strong scholarly attention to how knowledge is embedded in model training, transferred across domains, and operationalized through algorithmic mechanisms.

Cluster 2 constitutes the conceptual core of the AI–KM field, encompassing key nodes such as *artificial intelligence*, *knowledge management*, *machine learning*, *natural language processing*, and *knowledge sharing*. Its composition suggests that AI technologies have been broadly integrated into research on knowledge management, knowledge sharing, and semantic knowledge applications.

Cluster 3 extends toward privacy-preserving and intelligent system applications, represented by *federated learning*, *internet of things*, *blockchain*, and *data privacy*. This composition indicates that AI–KM research has also begun to **e**ngage with distributed intelligent systems and data governance concerns.

Taken together, these three clusters suggest that the AI–KM field is not organized around a single knowledge management theme. Rather, it spans technical learning mechanisms, AI-enabled knowledge management applications, and privacy-preserving and intelligent system applications.

Table 1: Major thematic clusters and top 10 keywords by occurrence in the AI–KM field

Cluster	theme label	Top 10 Author keywords by occurrence
Cluster 1	knowledge transfer and learning mechanisms	knowledge transfer (478); transfer learning (399); deep learning (385); convolutional neural network (154); knowledge distillation (132); neural network (115); domain adaptation (90); graph neural network (68); deep neural network (53); knowledge engineering (46)
Cluster 2	AI-enabled knowledge management applications	artificial intelligence (414); knowledge management (372); machine learning (332); natural language processing (134); knowledge sharing (81); large language model (68); knowledge graph (59); big data (50); ontology (49); data mining (48)
Cluster 3	Privacy-preserving and intelligent system applications	federated learning (75); internet of things (47); blockchain (23); edge computing (21); data privacy (18); digital twin (18); privacy (15); cloud computing (14); differential privacy (14); additive manufacturing (12)

3.2 Temporal Keyword Trend Analysis

To examine temporal shifts in research focus within the AI–KM field, this study adopts a cross-period keyword comparison to analyze changes in the frequency and structural composition of high-frequency keywords across different time intervals. The division of time periods should not be arbitrary, as the choice of period boundaries may affect the identification of emerging themes (Hoang, 2025). Given that the growth of AI–KM publications between 2015 and 2024 was uneven—showing marked acceleration after 2021—this study divides the observation window into three periods based on the turning points of the three-year moving average: 2015–2020 (P1), 2021–2022 (P2), and 2023–2024 (P3). To avoid interpretive bias arising from differences in corpus size across periods, Table 2 reports keyword prevalence as the proportion of documents containing each keyword relative to the total number of documents in the corresponding period.

As shown in Table 2, the research focus of the AI–KM field has evolved in clearly distinguishable stages. In P1, the most prominent keywords were *knowledge management* (18.9%), *knowledge transfer* (14.9%), *machine learning* (14.7%), *deep learning* (12.6%), and *artificial intelligence* (12.2%), indicating that research during this period was primarily built on the initial integration of established KM themes with AI-based techniques. In P2, *knowledge transfer* (18.8%) and *transfer learning* (15.5%) rose further, while *knowledge distillation* increased from 1.3% to 5.7%, reflecting a gradual shift from KM-oriented conceptual frameworks toward more technical topics centered on model transfer, knowledge compression, and representation learning. By P3, keywords such as *large language model* (5.1%), *ChatGPT* (2.1%), and *federated learning* (4.7%) had emerged rapidly, suggesting that generative AI and distributed learning have become prominent new research foci.

Building on these trends, several structural shifts merit closer discussion. The most analytically significant shift is the relative retreat of the *knowledge management* label within the field itself. As the defining term of this research domain, *knowledge management* declined from 18.9% in P1 to 8.8% in P3, a drop of 10.1 percentage points. This decline, however, does not necessarily indicate that KM research itself has receded. Rather, it can be interpreted as a shift in how researchers position their work: instead of labeling their studies directly under the broader KM umbrella, recent authors increasingly adopt more specific technical terms—such as *knowledge distillation*, *transfer learning*, and *knowledge graph*—to describe their research. In this sense, the intellectual center of gravity in AI–KM research appears to be shifting from "AI research framed around KM themes" toward "AI-based technical research that carries KM implications." This reading is consistent with

recent studies suggesting that AI is no longer merely an auxiliary tool layered onto KM processes, but has increasingly entered the core stages of knowledge creation, acquisition, sharing, and application, transforming how these activities are conducted (Jarrahi et al., 2023; Nakash & Bolisani, 2025).

Table 2: Top 20 author keywords across three developmental stages of the AI–KM field

Keyword	P1%	P2%	P3%	Δ(P2–P1)	Δ(P3–P2)	Δ(P3–P1)
knowledge transfer	14.9	18.8	15.1	+3.9	-3.7	+0.2
transfer learning	11.6	15.5	13.7	+3.9	-1.8	+2.1
deep learning	12.6	13.9	12.9	+1.3	-1.0	+0.3
artificial intelligence	12.2	14.0	12.5	+1.8	-1.5	+0.3
knowledge management	18.9	10.1	8.8	-8.8	-1.3	-10.1
machine learning	14.7	12.9	8.1	-1.8	-4.8	-6.6
knowledge distillation	1.3	5.7	6.7	+4.4	+1.0	+5.4
large language model	0.0	0.0	5.1	0.0	+5.1	+5.1
federated learning	0.4	2.3	4.7	+1.9	+2.4	+4.3
convolutional neural network	4.7	7.5	4.6	+2.8	-2.9	-0.1
graph neural network	0.5	2.7	3.7	+2.2	+1.0	+3.2
domain adaptation	3.2	2.7	3.6	-0.5	+0.9	+0.4
natural language processing	6.0	5.0	3.1	-1.0	-1.9	-2.9
knowledge sharing	2.2	2.4	3.0	+0.2	+0.6	+0.8
neural network	4.0	6.0	2.7	+2.0	-3.3	-1.3
knowledge graph	1.4	1.8	2.5	+0.4	+0.7	+1.1
fault diagnosis	0.6	2.0	2.4	+1.4	+0.4	+1.8
chatgpt	0.0	0.0	2.1	0.0	+2.1	+2.1
knowledge engineering	0.5	1.9	2.1	+1.4	+0.2	+1.6
contrastive learning	0.0	0.7	2.0	+0.7	+1.3	+2.0

Note. Percentages were calculated by dividing the number of documents containing a given keyword by the total number of documents in each period. The total number of documents was 855 in P1 (2015–2020), 735 in P2 (2021–2022), and 1,267 in P3 (2023–2024). Percentage-point differences are reported to facilitate cross-period comparison.

By contrast, *knowledge transfer* maintained a consistently high prevalence across all three periods (14.9%, 18.8%, and 15.1%), making it the most stable bridging term in the field and suggesting that questions of how knowledge is transferred, learned, and reused remain central to AI–KM research. Yet the thematic context in which this term appears has shifted noticeably over time. In P1, *knowledge transfer* occurred predominantly alongside KM-oriented terms centered on *knowledge management*; from P2 onward, it increasingly co-occurred with machine-learning-oriented terms such as *transfer learning*, *domain adaptation*, and *knowledge distillation*. This shift suggests that a single surface term may now carry two distinct research meanings: from a management perspective, the transfer of knowledge between organizations or individuals; from a machine-learning perspective, the transfer of knowledge between models. Drawing on the notion of interdisciplinary semantic drift (Wang et al., 2023), this phenomenon can be interpreted as a gradual relocation of a concept's semantic center across disciplinary communities.

Beyond these structural changes in established vocabulary, P3 also witnessed the rapid introduction of emerging technical topics. *Large language model* and *ChatGPT* rose from 0% in the earlier periods to 5.1% and 2.1%, respectively, and *federated learning* grew from 0.4% in P1 to 4.7% in P3. The emergence of these keywords does not appear to represent a linear extension of existing themes; rather, they were introduced in conjunction with a broader paradigm shift in AI, most notably the wave of generative AI that followed the public release of ChatGPT at the end of 2022. Notably, these new terms have not displaced established core concepts such as *knowledge transfer* and *deep learning* but instead coexist with them. This pattern suggests that recent developments in the AI–KM field may be better characterized as *paradigm layering* than *paradigm replacement*: new technologies are not establishing entirely separate research traditions but appear to be absorbed into existing problem spaces and reinterpreted within the field's prior knowledge context. Taken together, the relative retreat of traditional KM labels, the semantic shift of core bridging terms, and the rapid introduction of new technical topics constitute the three principal axes along which the research focus of AI–KM has been reorganized over the past decade.

4 Conclusions

This study examined how the knowledge structure of the AI–KM field has evolved amid its rapid expansion in recent years. Drawing on 2,857 AI–KM publications indexed in the Web of Science between 2015 and 2024,

the study combined descriptive analysis, co-word network analysis, and cross-period keyword comparison to conduct a longitudinal examination across three layers: publication trajectories, conceptual structure, and temporal changes in keywords. The findings suggest that the development of the AI–KM field has not followed a single linear path. Rather, it has been characterized by the relative retreat of the traditional knowledge management label, the semantic drift of *knowledge transfer* across disciplinary communities, and the rapid emergence of technical topics such as *large language model*, *ChatGPT*, and *federated learning*.

Taken together, these patterns are better characterized as paradigm layering than as paradigm replacement. Even as the field continues to expand in scale, its core problem orientations have not been displaced by new technologies. Instead, they have continued to evolve through shifts in vocabulary, the repositioning of meaning, and the layering of new topics onto existing research concerns.

The main contribution of this study lies in complementing existing bibliometric research, which has largely focused on cross-sectional description, by adopting a longitudinal perspective. In doing so, it extends understanding of the AI–KM field from a static snapshot toward an evolutionary view and provides an empirical foundation for future theory building, research positioning, and methodological development in this area.

5 Limitations and Future Directions

This study has several methodological limitations that should be acknowledged and addressed in future research.

First, the study relies on a single data source. Although the Web of Science Core Collection offers strong advantages in indexing quality and the completeness of citation data, it may exclude relevant conference papers that are indexed only in Scopus or IEEE Xplore. Future research could integrate multiple databases to improve coverage.

Second, the use of author keywords has inherent limitations. Although author keywords provide a direct reflection of how researchers position their work, their quality depends on individual labeling practices. Some publications do not provide keywords, while others use terms that are too generic to support fine-grained analysis. Future research could combine Keyword Plus, abstract-based keyword extraction, or topic modeling to reduce reliance on author-provided keywords alone.

Third, the observed shift in the semantic center of knowledge transfer across periods is primarily inferred from differences in co-occurring vocabulary. Future research could employ word embedding models to compare the semantic positioning of the same term across different periods, thereby providing a more rigorous way to examine the presence and magnitude of interdisciplinary semantic drift.

Future research should also integrate multiple databases, abstract-based corpora, and semantic analysis methods to more precisely trace conceptual drift and structural integration across periods, thereby deepening the study of knowledge evolution in the AI–KM field.

6 AI Declaration

In preparing this paper, the author, Hsuan-Hao Lo, used ChatGPT and Claude solely to assist with English language editing. As English is not the author's first language, these tools were used to improve grammar, refine wording, and enhance overall readability. The author reviewed and revised all AI-assisted outputs and takes full responsibility for the final content of the manuscript. All data were retrieved, cleaned, analysed, and interpreted by the author; no data and statistics presented in this paper were generated by AI tools.

References

Bernatović, I., Slavec Gomezel, A., & Černe, M. (2022). Mapping the knowledge-hiding field and its future prospects: a bibliometric co-citation, co-word, and coupling analysis. *Knowledge Management Research & Practice*, *20*(3), 394–409.

Cantu-Ortiz, F. J. (2021). Knowledge management and artificial intelligence analytics: A bibliometric study of research trends. In *A Research Agenda for Knowledge Management and Analytics* (pp. 67–88). Edward Elgar Publishing.

Cobo, M. J., López-Herrera, A. G., Herrera-Viedma, E., & Herrera, F. (2011). An approach for detecting, quantifying, and visualizing the evolution of a research field: A practical application to the Fuzzy Sets Theory field. *Journal of informetrics*, *5*(1), 146–166.

Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., & Lim, W. M. (2021). How to conduct a bibliometric analysis: An overview and guidelines. *Journal of business research*, *133*, 285–296.

Gaviria-Marin, M., Merigó, J. M., & Baier-Fuentes, H. (2019). Knowledge management: A global examination based on bibliometric analysis. *Technological Forecasting and Social Change*, *140*, 194–220.

Grant, M. J., & Booth, A. (2009). A typology of reviews: an analysis of 14 review types and associated methodologies. *Health information & libraries journal*, *26*(2), 91–108.

Hoang, A.-D. (2025). Evaluating bibliometrics reviews: A practical guide for peer review and critical reading. In (Vol. 49, pp. 1074–1102): SAGE Publications Sage CA: Los Angeles, CA.

Jarrahi, M. H., Askay, D., Eshraghi, A., & Smith, P. (2023). Artificial intelligence and knowledge management: A partnership between human and AI. *Business horizons*, *66*(1), 87–99.

Jiang, Z., Wang, Y., Li, W., Peng, K. Z., & Wu, C. H. (2023). Career proactivity: A bibliometric literature review and a future research agenda. *Applied Psychology*, *72*(1), 144–184.

Kushwaha, A., & Maurya, S. (2025). AI-Driven Knowledge Management: A Bibliometric Perspective and Future Directions. *Anveshak International Journal of Management (AIJM)*, *14*(2).

Marzi, G., Balzano, M., Caputo, A., & Pellegrini, M. M. (2025). Guidelines for bibliometric-systematic literature reviews: 10 steps to combine analysis, synthesis and theory development. *International Journal of Management Reviews*, *27*(1), 81–103.

Merigó, J. M., Mas-Tur, A., Roig-Tierno, N., & Ribeiro-Soriano, D. (2015). A bibliometric overview of the Journal of Business Research between 1973 and 2014. *Journal of business research*, *68*(12), 2645–2653.

Muñoz-Leiva, F., Viedma-del-Jesús, M. I., Sánchez-Fernández, J., & López-Herrera, A. G. (2012). An application of co-word analysis and bibliometric maps for detecting the most highlighting themes in the consumer behaviour research from a longitudinal perspective. *Quality & Quantity*, *46*(4), 1077–1095.

Nakash, M., & Bolisani, E. (2025). The transformative impact of AI on knowledge management processes. *Business Process Management Journal*, *31*(8), 124–147.

Öztürk, O., Kocaman, R., & Kanbach, D. K. (2024). How to design bibliometric research: an overview and a framework proposal. *Review of managerial science*, *18*(11), 3333–3361.

Tranfield, D., Denyer, D., & Smart, P. (2003). Towards a methodology for developing evidence-informed management knowledge by means of systematic review. *British journal of management*, *14*(3), 207–222.

Van Eck, N., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *scientometrics*, *84*(2), 523–538.

Wamba, S. F., Bawack, R. E., Guthrie, C., Queiroz, M. M., & Carillo, K. D. A. (2021). Are we preparing for a good AI society? A bibliometric review and research agenda. *Technological Forecasting and Social Change*, *164*, 120482.

Wang, Z., Peng, S., Chen, J., Kapasule, A. G., & Chen, H. (2023). Detecting interdisciplinary semantic drift for knowledge organization based on normal cloud model. *Journal of King Saud University-Computer and Information Sciences*, *35*(6), 101569.

Embedding ESG in Strategic Management Course: A System Dynamics Approach to Experiential Learning

MeiYi Winnie Song, Ju-We Chen
Department of Business Administration, Shih Chien University. Taipei, Taiwan

Abstract

This study aims to explore how to effectively integrate Environmental, Social, and Governance (ESG) issues into the strategic management curriculum and propose an experiential learning framework based on system dynamics (SD). Traditional strategic management courses may make it difficult for students to understand the feedback structures, time delays, nonlinear effects, and multi-objective trade-offs involved in ESG decision-making when they rely mainly on prescriptive concepts, case studies, or static analytical tools. Based on the methodology of design-based research, this study introduces causal loop diagrams, stock-and-flow diagrams, and scenario simulation activities into the undergraduate Strategic Management course, allowing students to test ESG strategy options in a simulated environment and observe the dynamic relationship between short-term financial performance, long-term environmental risks, social legitimacy, and governance mechanisms. Research data include pre- and post-course questionnaires, student reflection reports, model-building outcomes, and strategic decision-making performance. The preliminary descriptive findings indicate that experiential learning supported by system dynamics may help improve students' ESG literacy, systems thinking capability, and strategic foresight. Theoretically, this study considers ESG competencies as a dynamic knowledge capability formed through knowledge interaction, model building, and reflective practices. In terms of methods, it demonstrates the feasibility of system dynamics as a tool for management education research and curriculum design. In practice, it provides management educators with a replicable and scalable ESG strategy curriculum design framework.

Keywords: ESG Strategy; System Dynamics; Experiential Learning; Strategic Management Course; Knowledge-Based Sustainability

Paper type: Academic Research Paper

1 Introduction

As Environmental, Social, and Governance (ESG) issues have become central to corporate strategic decision-making, stakeholder management, and long-term value creation, business education must move beyond treating sustainability as a subsidiary topic of ethics or corporate social responsibility. Prior research argues that sustainability and ESG should be integrated into business and management curricula so that students can understand the relationships among sustainability strategies, organizational governance, and business performance (Beamond et al., 2026; Martínez-Bravo et al., 2024). This need is particularly evident in strategic management courses, where ESG should be framed not merely as a compliance requirement but as a strategic capability for risk identification, stakeholder response, competitive advantage, and organizational resilience. However, current ESG instruction often relies on conceptual explanation, case discussion, or indicator-based analysis, offering limited support for understanding the dynamic complexity of ESG decision-making. Sustainability education emphasizes that higher education should cultivate systems thinking, cross-domain integration, critical reflection, and future-oriented capabilities rather than only transmit static knowledge (Rieckmann, 2012; Svanström et al., 2008). To address this gap, this study proposes a system dynamics (SD)-based experiential learning approach for integrating ESG into strategic management curricula. System dynamics is suitable for analyzing feedback loops, time delays, nonlinear effects, and long-term mechanisms in sustainability issues, while experiential learning and business simulation can enhance student engagement, reflection, and attitude change in sustainable decision-making (Adib, 2024; Faham et al., 2017). Accordingly, this study aims to construct an integrated "ESG–Strategic Management–System Dynamics–Experiential Learning" curriculum framework, examine how causal loop diagrams (CLDs) and stock-and-flow diagrams (SFDs) influence students' ESG literacy, systems thinking, and strategic foresight, and explain how ESG capability can be transformed from static knowledge into actionable strategic knowledge capability. Its contributions are theoretical, methodological, and practical, offering a replicable ESG strategy curriculum design model for management education.

2 Literature Review

2.1 ESG and strategic management education

The core of ESG education is to enable students to understand that corporate sustainability is not merely associated with corporate ethics or social responsibility, but is interconnected with strategic choices, governance mechanisms, and long-term value creation. When introducing ESG into higher education, institutions should cultivate students' cross-disciplinary judgment and sustainability-oriented decision-making capabilities in the face of sustainability challenges through curriculum integration, research innovation, community engagement, leadership governance, and assessment feedback (Alenezi & Alanazi, 2024). In the context of business and management education, sustainability courses should not be limited to a single topic but should help students understand how companies transform ESG factors into strategic capabilities and organizational knowledge assets (Del Gesso & Parravicini, 2025). This also echoes the positioning of this study to integrate ESG into strategic management courses.

2.2 Experiential learning and sustainability ability development

Experiential learning provides an important teaching foundation for ESG education. Kolb (1984) argues that learning comes from a cyclical process of concrete experience, reflective observation, abstract conceptualization, and active experimentation. When applied to ESG strategy courses, students not only read corporate cases but also personally test the consequences of different sustainability strategies through simulations, team discussions, and reflection reports. Adib (2024) pointed out that business simulations incorporating sustainable decision-making can influence students' learning outcomes, satisfaction, and sustainability attitudes, indicating that experiential design helps translate ESG knowledge into decision-making understanding.

2.3 System dynamics and systems thinking

System dynamics provides a methodological basis for dealing with the complexity of ESG strategies. ESG decisions often involve feedback loops, time delays, and nonlinear effects, such as short-term cost pressures that may discourage environmental investment, but long-term environmental performance can affect brand trust and governance legitimacy. Faham et al. (2017) indicate that system dynamics can support the development of sustainability competencies in higher education by revealing the long-term behavior of sustainable systems through reinforcing and balancing feedback loops. Svanström et al. (2008) also argue that sustainable development learning outcomes should include systematic or holistic thinking, cross-perspective integration, critical thinking, and the ability to change action.

2.4 ESG capabilities from a knowledge management perspective

From the perspective of knowledge and ability, ESG literacy is not just conceptual knowledge but a dynamic ability formed by students in situational judgment, model construction, strategic simulation, and reflective practice. Rieckmann (2012) pointed out that future-oriented higher education should cultivate systematic thinking, foresight, normative, strategic and cooperative skills; Wiek et al. (2011) also consider these competencies as a core competency framework for sustainable education. Therefore, this study positions ESG education as a process of "knowledge capability generation", enabling students to transform abstract sustainability concepts into analytical, simulated, and applicable strategic management skills through experiential learning supported by system dynamics.

Existing studies have mostly discussed ESG education, experiential learning, or system dynamics separately, but rarely integrate the three in strategic management curriculum design. Therefore, this study proposes an integrated architecture of "ESG-Strategic Management-System Dynamics-Experiential Learning" (Figure 1).

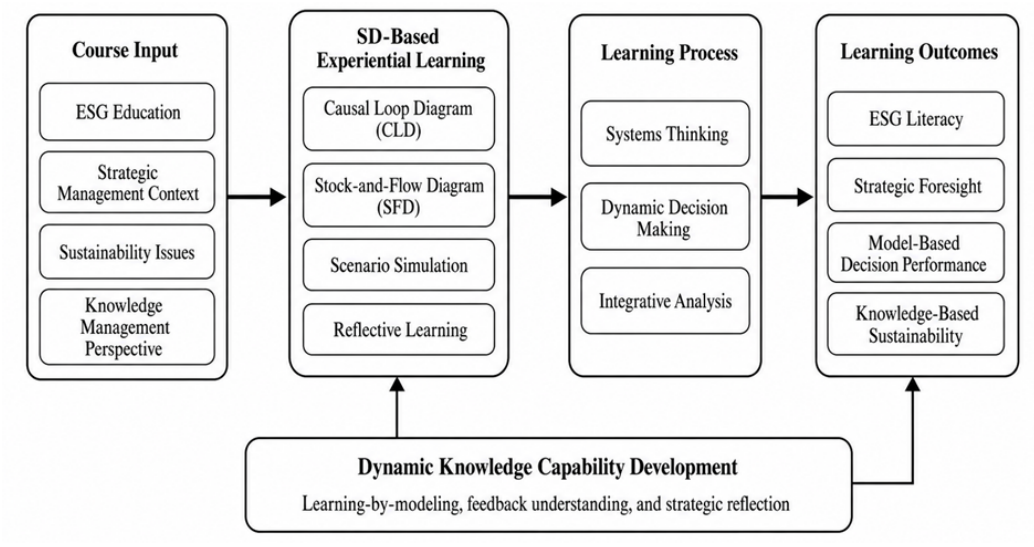


Figure 1. Research Framework of ESG-Embedded Strategic Management Learning

3 Research Methodology

This study adopts Design-Based Research (DBR) as its methodological basis. DBR connects educational theory and teaching practice through theory-informed curriculum design, teaching intervention, data collection, and iterative refinement in real educational settings. It is particularly suitable for examining innovative curricula, technology-supported learning, and complex competency development because it focuses not only on learning outcomes but also on how instructional design operates in authentic contexts (The Design-Based Research Collective, 2003; Wang & Hannafin, 2005; Anderson & Shattuck, 2012). Accordingly, this study applies DBR to examine how integrating ESG into a strategic management course through system dynamics-based experiential learning can enhance students' ESG literacy, systems thinking capability, and strategic foresight.

The research field is an undergraduate Strategic Management course. The curriculum transforms ESG issues from conceptual content into dynamic decision-making scenarios. The course intervention consists of four stages. First, students are introduced to ESG and strategic management concepts, including how ESG issues influence competitive advantage, stakeholder trust, and long-term value creation. Second, students receive system dynamics training, learning variables, causal relationships, reinforcing loops, balancing loops, time delays, system boundaries, causal loop diagrams, and stock-and-flow diagrams. Third, students construct ESG strategy models based on corporate or industry scenarios, incorporating variables such as environmental investment, carbon emissions, operating costs, brand trust, social legitimacy, governance transparency, and financial performance. Fourth, students conduct scenario simulation, decision-making, and reflection by testing alternative ESG strategies, observing long-term outcomes, and explaining decision logic, unintended consequences, and governance implications. This design aligns with the use of system dynamics for complex system modeling, policy simulation, and strategic learning (Sterman, 2000).

3.1 Data Collection

Data collection adopted a mixed-methods design to enhance the completeness and credibility of the research findings. The study was implemented in a 16-week Strategic Management course for third-year students in the Department of Business Administration, with 42 students participating. Quantitative data were collected through pre- and post-course questionnaires designed by the researchers based on relevant literature. The questionnaire measured changes in students' ESG literacy, systems thinking capability, strategic foresight, and model-based decision-making capability. Qualitative data included student reflection reports, group system dynamics model artifacts, class discussion records, and teacher observation notes. The student model artifacts were evaluated independently by two researchers using the evaluation rubric. Any scoring differences were discussed and resolved through consensus to improve the reliability of the assessment process. Faham et al. (2017) indicated that system dynamics research in sustainability education can integrate multiple data sources, including literature, questionnaires, interviews, and observations, to construct and validate sustainability-related learning models. This supports the use of mixed methods for examining complex learning processes in sustainability and ESG education.

Table 1: Mixed-method Data Sources and Descriptions

Sources	Description
Pre-course/post-course questionnaires	Evaluate changes in ESG literacy, systems thinking ability, and strategic foresight ability
Student model artifacts	Analyze the completeness and logic of students' construction of CLDs / SFDs
Simulation decision records	Examine whether students can balance short-term and long-term ESG strategy outcomes
Reflection reports	Analyze students' understanding of unintended consequences, stakeholders, and governance logic
Teacher observation records	Supplement student interaction, discussion, and learning difficulties

3.2 System dynamics model architecture

The ESG strategy learning model designed in this study includes the following variables (Table 2):

Table 2: Core Variables of ESG System Dynamics Learning Model

Variable Type	Variables
ESG input variables	Environmental Investment, Social Engagement, Governance Mechanisms
Strategy process variables	Stakeholder Trust, Operational Efficiency, ESG Integration Capability
Outcome variables	Financial Performance, Environmental Impact, Social Legitimacy, Strategic Resilience
Learning variables	ESG Literacy, Systems Thinking Capability, Strategic Foresight

The causal loop diagram (Figure 2) includes three reinforcing loops and two balancing loops: R1: ESG knowledge learning - strategy integration cycle, R2: stakeholder trust - social legitimacy cycle, R3: simulation reflection - strategic foresight cycle, B1: short-term cost pressure-ESG investment constraint cycle and B2: cognitive load - learning effectiveness regulation loop.

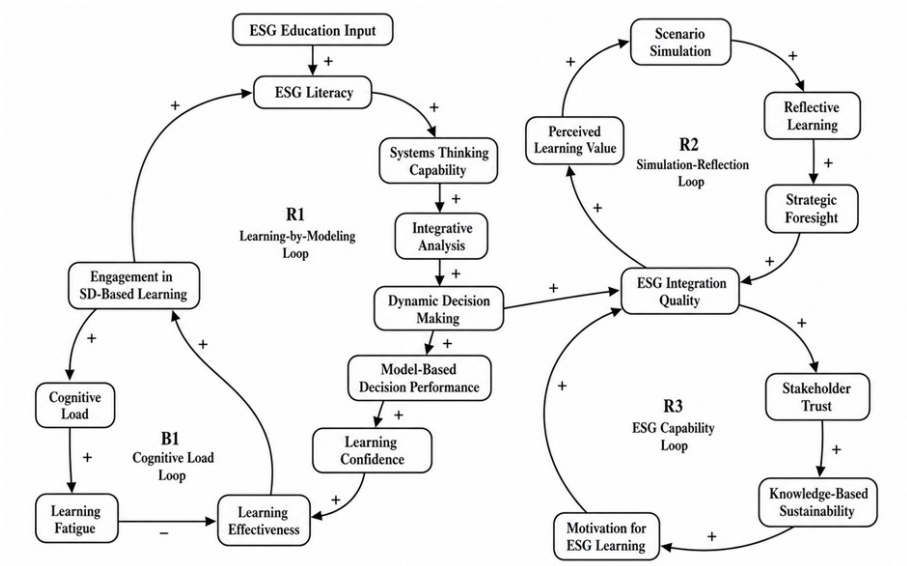


Figure 2. Causal Loop Diagram of ESG Strategic Learning System

The stock-and-flow diagram (Figure 3) shows that the main stocks are ESG Literacy, Systems Thinking Capability, Strategic Foresight, and Decision Performance.

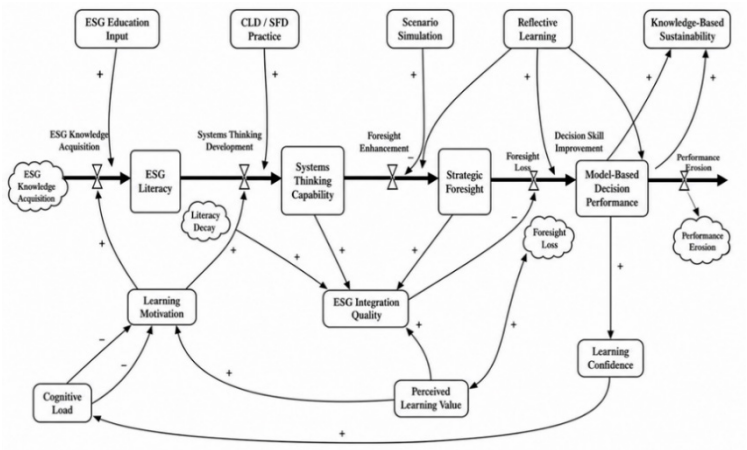


Figure 3. Stock-and-Flow Model of ESG Competence Development

3.3 Data analysis methods

In terms of analysis methods, quantitative data were first analyzed using descriptive statistics to compare pre- and post-course learning outcomes. Paired-sample tests, or the Wilcoxon signed-rank test if the data were not normally distributed, will be used in subsequent inferential analysis. Qualitative data were analyzed using thematic analysis following Braun and Clarke's (2006) procedure. Student-developed model artifacts were evaluated using scoring criteria covering variable completeness, causal logic, feedback loops, time delays, and strategic explanatory capability. This consensus-based procedure was used to enhance interpretive reliability in evaluating student-developed system dynamics models.

3.4 Summary

This research method emphasizes the cyclical process of "design–implementation–evaluation–revision" and examines how system dynamics supports ESG strategy learning through curriculum intervention and multiple data sources. This approach not only responds to the needs of sustainable competency development in management education but also helps establish a curriculum model for ESG strategy management that can be replicated, adjusted, and sustainably improved.

4 Research Results

4.1 Improvement of ESG literacy and strengthening of systems thinking capability

Based on the preliminary analysis of curriculum implementation data, the results of this study can be summarized into three aspects: learning effectiveness, model outcomes, and reflective analysis. First, in terms of learning effectiveness, students' understanding of ESG appeared to shift from "corporate social responsibility (CSR) or compliance requirements" to "strategic decision-making and long-term value creation capabilities" after the course. This result is consistent with the claim that Education for Sustainable Development (ESD) emphasizes systems thinking, cross-perspective integration, critical thinking, and change action (Svanström et al., 2008). Before class, students often use single concepts such as environmental protection, social responsibility, or governance disclosure to understand ESG. After the course, students could further explain the relationship between ESG and competitive advantage, stakeholder trust, risk governance, and financial performance. This change also echoes the emphasis on systematic, strategic, and integrated learning outcomes in sustainability assessment research (Redman et al., 2021).

4.2 Model construction results

Table 3 lists the evaluation criteria used to evaluate the system dynamics model developed by students. In addition to technical modeling quality, this evaluation criterion also evaluates ESG integration, strategy interpretation, scenario simulation, and reflective learning.

Table 3. Evaluation Rubric for Student System Dynamics Models

Construct	Beginning	Developing	Proficient	Advanced
Variable identification	Variables are vague, incomplete, or only loosely linked to ESG and strategic management.	Some ESG and strategic variables are identified, but important variables are still missing.	Key environmental, social, governance, and strategic performance variables are clearly identified.	The variables are complete, theoretically based, and highly linked to ESG strategies, stakeholder governance, and long-term value creation.
Causal logic	Causal links are unclear, inconsistent, or lack reasonable explanations.	Basic causal relationships are presented, but some links lack explanation or logical consistency.	Causal relationships are generally clear, consistent, and reasonably explained.	Causal logic is rigorous and coherent, supported by ESG, strategic management, or system dynamics perspectives.
Feedback loop structure	No feedback loops are created, or the loops are incorrectly constructed.	Some reinforcing loops or balancing loops are included, but the loop logic is not yet complete.	Reinforcing and balancing loops are correctly identified and labeled.	Feedback loops can clearly explain the dynamics of ESG, including the relationship between learning, trust, cost pressure, governance, and strategic foresight.
Stock-and-flow representation	Confusing stocks and flows, or failing to correctly distinguish between the two.	Basic stocks and flows are included, but the relationship is not yet complete.	Stocks, inflows, outflows, and auxiliary variables are appropriately presented.	Models can clearly represent the dynamic process of capability accumulation, learning decline, feedback effects, and changes over time.
Time delay and dynamic behavior	Time delay and long-term effects are not considered.	Some delay effects are mentioned, but not yet clearly integrated into the model.	Time delays and long-term outcomes are presented in key relationships.	Models can effectively explain short-term and long-term trade-offs, unintended consequences, and ESG learning trajectories.
ESG integration	ESG is only considered a general concept or a standalone issue.	ESG factors are included, but are weakly linked to strategic management.	ESG dimensions are integrated with strategic management decisions and learning outcomes.	ESG is fully embedded as a dynamic strategy system that connects governance, stakeholder trust, sustainability performance, and knowledge capabilities.
Strategic interpretation ability	The model provides only limited strategic implications.	Some strategic implications are presented, but the analysis remains descriptive.	Models can support sound ESG strategic decision interpretation.	Models generate deep strategic insights across resource allocation, stakeholder management, governance mechanisms, and long-term value creation.
Scenario simulation quality	Scenario assumptions are absent or unclear.	Scenario contexts are included, but their scope is limited and weakly connected to the model.	Scenario design is logical and can be linked to model variables.	Scenario simulations can effectively compare different ESG strategies and account for differences in long-term outcomes.
Reflective learning	Reflection is superficial or lacks relevance to model results.	Learning experiences are described, but analytical insights remain limited.	How model results changed students' understanding of ESG strategies is explained.	Reflection demonstrates deep learning, critical thinking, unintended consequences recognition, and strategic hypothesis modification skills.
Overall model consistency	The model is fragmented and difficult to understand or interpret.	The model has a basic structure, but the overall consistency is insufficient.	The model is consistent, readable, and aligned with the curriculum objectives.	The model is theoretically grounded, visually clear, analytically practical, and suitable for supporting ESG strategic decision-making.

Note. The rubric may be scored on a 1–4 scale: 1 = Beginning, 2 = Developing, 3 = Proficient, and 4 = Advanced. Total scores can be used to evaluate students' system dynamics model quality, while individual dimensions can help diagnose specific learning strengths and weaknesses.

In terms of model results, students can transform ESG strategic issues into a dynamic system with a feedback structure through causal loop diagrams and stock-and-flow diagrams. Most teams can identify the reinforcing loop of "ESG knowledge learning-strategic integration ability-decision-making quality-learning confidence" and the balancing loop of “short-term cost pressure–ESG investment willingness–sustainable performance improvement.” This indicates that system dynamics tools helped students understand time delays, nonlinear variations, and unintended consequences in sustainability issues. Faham et al. (2017) pointed out that system dynamics can help higher education identify reinforcing and balancing mechanisms in sustainability development and understand long-term behavior changes through simulation. The student-developed models in this study show similar characteristics.

4.3 Strategic Decision-Making and Reflective Performance

Table 4 summarizes the comparison of students' learning outcomes before and after taking the ESG-embedded strategic management course. The preliminary descriptive results indicate positive changes across all learning dimensions. Among them, the improvement in model-based decision-making ability is the largest descriptive improvement, indicating that the use of causal loop diagrams, stock-and-flow models, and scenario simulations helps students translate ESG knowledge into actionable strategic analysis. Additionally, improved systems thinking capability and strategic foresight indicate that students are better able to identify feedback relationships, time delays, and long-term impacts embedded in ESG-related strategic decisions. Overall, the preliminary findings suggest the potential effectiveness of system dynamics-based experiential learning in enhancing ESG literacy, integrative analytical capability, and knowledge-based sustainability competence.

Table 4. Pre- and Post-Course Learning Outcome Comparison

Dimension	Measurement focus	Pre-course mean	Post-course mean	Change	Interpretation of the results
ESG literacy	Understanding of ESG concepts, strategic relevance, and sustainability implications	3.12	4.21	+1.09	Students can better understand that ESG is a strategic management issue, not just a compliance or corporate social responsibility issue.
Systems thinking capability	The ability to identify feedback loops, time delays, causal relationships, and systemic consequences	2.95	4.08	+1.13	When students analyze ESG issues, they are better equipped to grasp their dynamics and interconnectedness.
Strategic foresight	The ability to predict long-term consequences, unintended outcomes, and stakeholder impacts	2.88	4.05	+1.17	Students are better able to evaluate short-term strategic decisions from a long-term sustainability perspective.
Model-based decision performance	The ability to use causal loop diagrams and stock-and-flow diagrams to support strategic decision-making	2.76	3.98	+1.22	Students improved in applying system dynamics models to analyze ESG strategy choices.
Integrative analytical capability	The ability to connect financial, environmental, social, and governance factors	3.01	4.14	+1.13	Students can more effectively integrate ESG aspects into strategic analysis.
Reflective learning ability	The ability to reflect on simulation results and revise strategic assumptions	3.18	4.26	+1.08	Students are more aware of the feedback effects, strategic trade-offs, and governance logic in ESG decision-making.
Knowledge-based sustainability competence	The ability to translate ESG knowledge into actionable strategic knowledge	2.91	4.10	+1.19	Through model construction and reflection, students gradually form ESG competence as a dynamic knowledge capability.

Note. Measurement scale: 1 = strongly disagree; 5 = strongly agree. These values represent preliminary descriptive results; inferential statistical testing should be conducted in subsequent analysis.

In terms of reflective analysis, the student reflection reports show that experiential learning prompts students to shift from passively accepting ESG concepts to actively examining the long-term consequences of

strategic choices. Some students pointed out that if companies only pursue short-term profits, they may reduce environmental investment and governance improvements, resulting in reputational risks and a decline in stakeholder trust. Conversely, students also recognized that stronger governance transparency and resource allocation mechanisms could support long-term strategic resilience. This is consistent with experiential learning and business simulation research. Adib (2024) indicates that business simulations incorporating sustainable decision-making can affect students' learning outcomes, satisfaction, and sustainable attitudes, and a recent review of sustainable experiential learning also pointed out that such teaching can promote the formation of students' knowledge, abilities, values, and action (Sundman et al., 2025).

5 Discussion

5.1 Integration of Preliminary Findings

Overall, three preliminary findings can be identified. First, system dynamics appears to help transform ESG from abstract concepts into visual and actionable strategic issues. Second, experiential simulations appear to support students in understanding time delays, feedback effects, and unintended consequences in ESG decision-making. Third, model building and reflective activities appear to facilitate the transformation of ESG knowledge from conceptual understanding into strategic judgment capability. These findings suggest that system dynamics-based experiential learning may provide a useful pedagogical approach for integrating ESG into strategic management education.

Taken together, these findings indicate that system dynamics-supported experiential learning helps reposition ESG as a decision-making system that can be modeled, simulated, and reflected upon in strategic management courses. Students gradually develop ESG literacy, systems thinking capability, and strategic foresight capability during the learning process, and can transform sustainability concepts into knowledge and capabilities with governance logic and strategic implications. These findings are also in line with the original design of this study: through CLDs, SFDs, and simulation-based learning modules, students can understand the dynamic relationships among financial performance, environmental impact, social legitimacy, and governance mechanisms.

5.2 Implications

This study offers three implications. Theoretically, it repositions ESG education from “knowledge transfer” to “knowledge capability development.” ESG literacy is not limited to understanding ESG concepts; it also involves the ability to identify dynamic relationships among stakeholder interests, long-term risks, governance responsibilities, and value creation in strategic contexts. This aligns with sustainability education research, which argues that higher education should cultivate systems thinking, foresight, normative judgment, strategic action, and collaboration competencies (Wiek et al., 2011). Thus, ESG competence is conceptualized as a dynamic knowledge capability developed through curriculum design, model building, and reflective practice.

Methodologically, the study demonstrates that system dynamics is useful for addressing complex sustainability issues in management education. Unlike traditional ESG teaching based mainly on case reading or indicator analysis, causal loop diagrams and stock-and-flow diagrams help students visualize feedback loops, time delays, and nonlinear changes. This supports prior research showing that system dynamics can develop sustainability competencies and reveal long-term system behavior through reinforcing and balancing loops (Faham et al., 2017).

Practically, the study provides a replicable ESG strategy curriculum model. Through Design-Based Research, instructors can apply a “design–implementation–evaluation–revision” cycle to connect theory, instructional design, and classroom practice (The Design-Based Research Collective, 2003). This study further shows how sustainability-related knowledge can be structured, simulated, and transformed into actionable strategic knowledge through system dynamics-based learning.

6 Conclusion

This study proposes a curriculum design framework for integrating ESG issues into strategic management courses through experiential learning and system dynamics. The findings suggest that ESG education should move beyond concept delivery, case reading, or indicator-based analysis, because such approaches may be insufficient for cultivating the systems thinking and decision-making capabilities required to address complex sustainability challenges. By using causal loop diagrams, stock-and-flow diagrams, and scenario simulations, students can transform ESG issues into dynamic strategic systems involving feedback relationships, time delays,

trade-offs, and long-term consequences. This approach echoes sustainability education research that emphasizes systems thinking, strategic competence, and integrated problem-solving (Wiek et al., 2011), and supports the view that system dynamics can cultivate sustainability competencies in higher education (Faham et al., 2017).

The study makes three main contributions. Theoretically, it conceptualizes ESG literacy as a dynamic knowledge capability rather than static knowledge content. Through model construction, simulation-based decision-making, and reflective practice, students gradually develop sustainability-oriented judgment that can be applied to strategic contexts. Methodologically, the study demonstrates that design-based research is a useful approach for management education innovation, allowing curriculum design to respond to both theoretical development and practical teaching needs through a "design–implementation–evaluation–revision" cycle (Anderson & Shattuck, 2012). Practically, it provides management educators with a replicable ESG strategy curriculum model, positioning ESG as a core component of strategic management rather than a supplementary topic.

This study also has limitations. First, the research was conducted in a single course, and the sample size limits the generalizability of the findings. Second, learning outcomes were mainly assessed through pre- and post-course questionnaires, model artifacts, and reflection reports, making it difficult to evaluate students' long-term application of ESG capabilities in real organizational settings. Third, students' varying familiarity with system dynamics may have influenced model quality and learning performance. Future research may extend this framework to different institutions, course levels, and cross-national teaching contexts, and adopt longitudinal designs to examine the sustained development of ESG strategic capabilities. Further studies may also incorporate artificial intelligence (AI)-assisted modeling, learning analytics, and industry-based projects to explore how intelligent knowledge tools support sustainability-oriented management education.

7 Ethics Statement

This study involved course-based educational data. Students were informed of the research purpose, participation was voluntary, and non-participation did not affect course grades. All learning data were anonymized before analysis.

8 Artificial Intelligence Declaration

The authors used ChatGPT to support language editing, citation checking, and argument refinement. The authors reviewed, revised, and take full responsibility for the final content.

References

Adib, H. (2024). Experiential learning in higher education: Assessing the role of business simulations in shaping student attitudes towards sustainability. *The International Journal of Management Education*, *22(2)*, 100968.

Alenezi, M., & Alanazi, F. (2024). Integrating environmental social and governance values into higher education curriculum. *International Journal of Evaluation and Research in Education*, *13(5)*, 3493–3503.

Anderson, T., & Shattuck, J. (2012). Design-based research: A decade of progress in education research? *Educational Researcher*, *41(1)*, 16–25.

Beamond, M. T., Schmitz, M., Cordova, M., Ilieva, M. V., Zhao, S., & Panina, D. (2026). Sustainability in business education: A systematic review and future research agenda. *Critical Perspectives on International Business*, *22(1)*, 56–83.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3(2)*, 77–101.

Del Gesso, C., & Parravicini, P. (2025). Sustainability-focused business curricula educating corporate sustainability professionals: A knowledge-based view from Italian universities. *Journal of Innovation & Knowledge*, *10(5)*, 100771.

Faham, E., Rezvanfar, A., Mohammadi, S. H. M., & Nohooji, M. R. (2017). Using system dynamics to develop education for sustainable development in higher education with the emphasis on the sustainability competencies of students. *Technological Forecasting and Social Change*, *123*, 307–326.

Kolb, D. A. (1984). *Experiential Learning: Experience as the Source of Learning and Development*. Englewood Cliffs, NJ: Prentice Hall.

Martínez-Bravo, M. M., Capobianco-Uriarte, M. M., Terán-Yépez, E., Marín-Carrillo, G. M., & Casado-Belmonte, M. P. (2024). Integrating sustainability into business and management studies in higher education. The International Journal of Management Education, *22(1)*, 100939.

Redman, A., Wiek, A., & Barth, M. (2021). Current practice of assessing students' sustainability competencies: A review of tools. *Sustainability Science*, *16(1)*, 117–135.

Rieckmann, M. (2012). Future-oriented higher education: Which key competencies should be fostered through university teaching and learning? *Futures*, *44(2)*, 127–135.

Sterman, J. D. (2000). *Business dynamics: Systems thinking and modeling for a complex world*. Irwin/McGraw-Hill.

Sundman, J., Feng, X., Shrestha, A., Johri, A., Varis, O., & Taka, M. (2025). Experiential learning for sustainability: A systematic review and research agenda for engineering education. *European Journal of Engineering Education*, 1–31.

Svanström, M., Lozano-García, F. J., & Rowe, D. (2008). Learning outcomes for sustainable development in higher education. *International Journal of Sustainability in Higher Education*, *9(3)*, 339–351.

The Design-Based Research Collective. (2003). Design-based research: An emerging paradigm for educational inquiry. *Educational Researcher*, *32(1)*, 5–8.

Wang, F., & Hannafin, M. J. (2005). Design-based research and technology-enhanced learning environments. *Educational Technology Research and Development*, *53(4)*, 5–23.

Wiek, A., Withycombe, L., & Redman, C. L. (2011). Key competencies in sustainability: A reference framework for academic program development. *Sustainability Science*, *6(2)*, 203–218.

A System Dynamics Analysis of Career Agency Formation among Gen Z Interns: Unpacking the Contextual Mechanisms of Overseas Work-Integrated Learning

MeiYi Winnie Song , Kuan-Lun Chen
Department of Business Administration, Shih Chien University. Taipei, Taiwan

Abstract

This study explores how career agency is formed in multiple relationship contexts when Generation Z (Gen Z) students participate in overseas internships. In the face of artificial intelligence (AI)-driven digital transformation and global labor market uncertainty, Generation Z's career development has shifted from a traditional linear employment path to a non-linear journey that emphasizes independent exploration, cross-cultural adaptation, and proactive opportunity building. Based on Work-Integrated Learning (WIL) and career agency literature, this study adopts a System Dynamics (SD) approach to construct causal loop diagrams (CLDs) and a stock-and-flow model to analyze the feedback relationships among university support, corporate mentoring, peer networks, cross-cultural pressure, career self-management, social capital, and employability capital. The proposed model suggests that career agency is not a static psychological trait, but a dynamic capability accumulated through reinforcing and balancing loops across pre-internship preparation, internship interaction, and post-internship transition. Scenario simulations indicate that program designs that integrate university support, corporate mentoring, and peer networks can best promote the simultaneous growth of career agency, social capital, and employability capital. This study contributes by connecting career development theory, WIL research, and System Dynamics (SD), while providing managerial implications for higher education institutions designing overseas internship and youth career development programs.

Keywords: Generation Z, overseas internships, career agency, Work-Integrated Learning, System Dynamics

Paper type: Academic Research Paper

1 Introduction

In the context of artificial intelligence (AI)-driven digital transformation and high uncertainty in the global labor market, the career development of Generation Z (Gen Z) college students has gradually shifted from the traditional linear employment path to a non-linear journey that emphasizes independent exploration, cross-domain mobility, and active construction opportunities. Schroth (2019) suggested that Generation Z has shown different work values and behavior patterns after entering the workplace, and organizations and educational institutions must re-understand their career expectations. Work-Integrated Learning (WIL) is regarded as an important mechanism to enhance graduates' employability, especially overseas internships can enhance students' career readiness through cross-cultural work experience, industry contact, and social networking (Jackson & Dean, 2023).

However, existing overseas internship research mostly focuses on outcome variables such as employability, cross-cultural ability, or career self-efficacy, and rarely explores how these abilities are gradually generated in the context of pre-internship preparation, internship interaction, and post-internship transfer. International students often face challenges such as language, culture, workplace norms, and insufficient support resources when participating in WIL (Vu et al., 2022), indicating that the effectiveness of overseas internships does not depend solely on individual efforts but is influenced by factors such as university systems, corporate mentoring, peer support, and cross-cultural adaptation. On the other hand, career agency and career self-management skills have been regarded as key capabilities in the face of labor market uncertainty. Relevant studies have also suggested that career competencies can be translated into career success and movement capital through job crafting, career engagement, and educational support (Akkermans & Tims, 2017; Petruzziello et al., 2022).

Based on this, this study takes the "Generation Z Overseas Internship Relationship Context" as the analysis object, adopts a System Dynamics (SD) perspective, and regards career agency as a dynamic ability accumulated through career self-management, social capital, employability capital, and institutional support in the multi-stage overseas internship process, rather than a single static psychological trait. This study therefore applies causal loop diagrams (CLDs) and stock-and-flow diagrams (SFDs) to analyze the dynamic formation of career agency in overseas WIL contexts. The research questions are as follows: (1) What relational factors affect the formation of career agency among Gen Z students during overseas internships? (2) What causal feedback relationships exist among university support, corporate mentoring, social capital, and career self-management?

(3) How do different internship support designs influence students' long-term employability and career resilience through SD mechanisms? The purpose of this study is to connect career development theory, WIL research, and System Dynamics (SD), to make up for the shortcomings of the existing literature on linear causal analysis, and to provide a dynamic decision-making basis for the design of overseas internship programs (Richardson, 2020; Patton & McMahon, 2006).

2 Literature review

2.1 Generation Z's career characteristics and work values

Growing up in a digital, platform economy, and highly uncertain labor market, Generation Z's career development has shifted from within traditional organizations to independent exploration, flexible mobility, and personalized career construction. Schroth (2019) suggested that after entering the workplace, Generation Z shows the characteristics of valuing feedback, development opportunities, work meaning, and psychological safety, and companies and higher education institutions need to re-understand their work values. Drewery et al. (2023) further found that Gen Z students' WIL experiences affect their emphasis on employer attributes and job value, indicating that internships are not just skills training but an important process in the formation of career cognition and career choice preferences.

2.2 Overseas internship and WIL

WIL is widely regarded as an important educational design to enhance the employability of college students, especially internships, workplace practice, and industry-academia cooperation can help students transform academic knowledge into workplace abilities. Jackson and Dean (2023) suggested that different types of WIL can contribute to graduate employability, but their effectiveness will be affected by program design, student participation, and the quality of workplace support. For overseas internships, WIL is more cross-cultural learning, and students must face language, life adaptation, organizational norms, and professional task challenges at the same time. A literature review by Vu et al. (2022) suggested that international students often encounter challenges such as cultural barriers, difficulties in workplace integration, and insufficient institutional support in WIL, indicating that the effectiveness of overseas internships cannot be measured solely by whether they have internship opportunities, but must be integrated into the relationship between schools, companies, peers, and cross-cultural environments. Overseas internships, as an international form of WIL, serve three functions: (1) a learning function, through which classroom knowledge is translated into practical workplace skills; (2) a career exploration function, through which students understand industries, roles, and person–career fit; and (3) a social capital function, through which employment opportunities and career resources are developed through supervisors, peers, and cross-cultural networks.

2.3 Career Agency and Career Self-Management

Career agency can be understood as an individual's ability to actively plan, adjust, and shape their career path. Akkermans and Tims (2017) suggested that career competencies can influence career success through job crafting, indicating that individuals do not passively accept the career environment but can create development opportunities through actions. Career ownership has been identified as an important factor shaping students' perceived employability (Donald et al., 2019). Petruzziello et al. (2022) also suggested that students' career investment and movement capital not only depend on personal initiative but are also influenced by the support of teaching staff. Wilhelm et al. (2024) regard career self-management behaviors as the foundation of multifaceted employability, including information search, goal setting, network building, and career monitoring. It can be seen that career agency is not a fixed personality trait, but a dynamic ability that gradually accumulates through educational support, workplace interactions, and personal reflection (Pham et al., 2023). The core concept of career agency is summarized in Table 1 of the role of this study.

Table 1. Core Concepts of Career Agency

Concept	Definition	Role in This Study
Career agency	The ability to proactively plan and shape your career	Core stock/outcome variables
Career self-management	Search for information, set goals, take action	Key process in career agency formation
Social capital	Resources formed by mentors, supervisors, peers, and cross-cultural networks	Affect employability and career confidence

Concept	Definition	Role in This Study
Employability capital	Skills, experience, resume signals, and workplace adaptability	The cumulative basis of career agency
Cross-cultural adaptation	The ability to adapt to differences in language, culture, and work norms	may promote or inhibit the effectiveness of the internship

2.4 System Dynamics (SD) and Career Development Research

A systems perspective helps compensate for the limitations of existing career and WIL research, which often relies on linear causal analysis. Patton and McMahon (2006) proposed a systematic theoretical framework for career development, emphasizing the interaction between individual, social, and environmental factors. Sterman (2002) suggested that System Dynamics (SD) is suitable for analyzing complex social systems with feedback, delay, and nonlinear characteristics. Pavlov et al. (2025) also explain that System Dynamics (SD) can understand the complex causal relationships in social and economic systems through feedback thinking. Based on this, this study treats Generation Z's overseas internships as a dynamic system, and analyzes how university support, corporate mentoring, peer networks, cross-cultural pressure, career self-management, social capital, and employability capital jointly affect the formation of career agency through causal loop diagrams and stock-and-flow models. The core idea of this study is to analyze the formation mechanism of Gen Z overseas WIL and career agency using the SD method.

3 Research methodology

3.1 Study design

This study uses System Dynamics (SD) as the main research method to analyze how career agency is gradually formed in multiple relationship contexts during the overseas internship process of Generation Z. SD is suitable for dealing with social systems with feedback, delay, cumulative, and nonlinear characteristics; Forrester (1994) suggested that System Dynamics (SD) can be used to understand and improve complex system structures, and Sterman (2000, 2002) also emphasized that model construction should combine problem definition, dynamic hypotheses, model formulation, testing, and policy simulation. Therefore, this study does not regard overseas internship as a single educational activity, but as a dynamic development system of "pre-internship preparation-internship interaction-post-internship transfer". The structure of this study design is summarized in Figure 1.

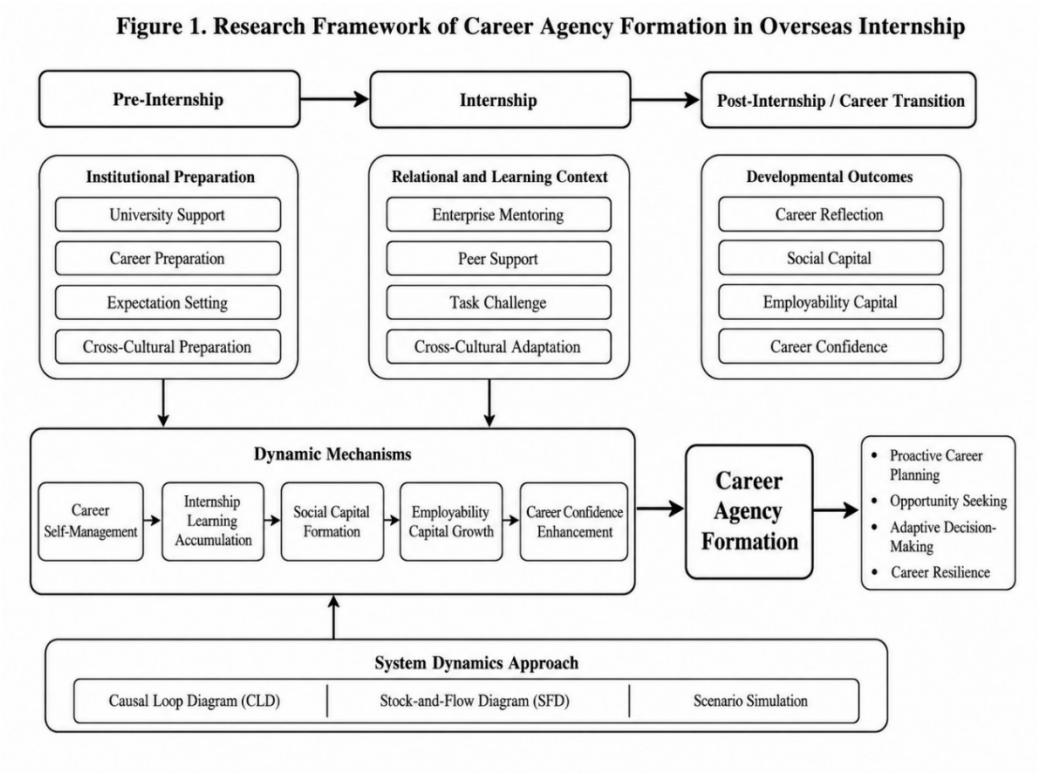


Figure 1. Research Framework of Career Agency Formation in Overseas Internship

3.2 System boundaries and analysis stages

In terms of modeling procedures, this study first extracts key variables based on WIL, career agency, and career self-management literature, and then uses a causal loop diagram (CLD) to present the positive and negative relationships between variables, and finally converts them into stock-and-flow diagrams (SFDs) to simulate the impact of different overseas internship designs on career agency formation. This study is set up to analyze the cumulative relationship between career agency, social capital and employability capital using CLDs and SFDs, and to identify leverage points for scenario simulation identification schemes. This study sets the system boundary into three phases (Table 2).

Table 2: System boundary stages

stage	Time frame	Main mechanism
Pre-internship stage	before and after application, training, and matchmaking	Career expectations, university support, language and cultural preparation
Internship stage	During overseas work and life	Corporate mentoring, task challenges, cross-cultural adaptation, peer support
Post-internship stage	Returning to the home country, employment, or entering a career program	Career reflection, resume transformation, employability accumulation, career decision-making

3.3 Variable settings

The core stock variables of this study include three items: first, career agency stock, which refers to students' cumulative ability to actively set career goals, seek resources, and adjust career paths; second, employability capital stock, including professional skills, cross-cultural abilities, work experience and resume signals; Third, social capital stock refers to the relational resources formed by university faculty, corporate mentoring, peers, and overseas workplace networks. The main flow variables include career self-management behavior, internship learning accumulation, social network formation, and career reflection. The auxiliary variables include university support quality, corporate mentor feedback, task challenge level, cross-cultural pressure, peer support, and career confidence (Table 3). This variable design echoes the view that WIL can improve employability, but its effect is affected by the type of internship, the quality of participation, and the support conditions. At the same time, career self-management behavior is also regarded as an important mechanism for maintaining and enhancing employability (Lasrado et al., 2024; Wilhelm et al., 2024).

Table 3: Core variables

Type	Variable name	Description
Stock variables	Career agency	Students' ability to actively shape career directions and actions
Stock variables	Employability Capital	Work skills, workplace experience, cross-cultural competencies, and resume signals
Stock variables	Social capital	Mentors, supervisors, peers, and workplace networking resources
Flow variables	Career self-management behavior	Goal setting, information search, career planning, and proactive action
Flow variables	Internship learning accumulation	task learning, problem solving, reflection and ability transformation
Auxiliary variables	University support quality	Pre-class training, matchmaking, counseling and crisis management
Auxiliary variables	Corporate guidance quality	Supervisor feedback, task design, and workplace support
Auxiliary variables	Cross-cultural pressure	Stress caused by differences in language, life, culture and work norms
Auxiliary variables	Career confidence	subjective beliefs about future career choices and employability

3.4 Modelling program

This study develops a literature-based exploratory SD model. Literature data are used to establish the initial variables and causal relationships. Future empirical data, including student surveys, interviews, internship reflection records, administrative data, and expert review, will be used to calibrate and validate the model.

The modelling process is as follows: 1. Key variables affecting career agency by literature extraction. 2. Establish positive and negative causal relationships between variables. 3. Draw a causal loop diagram to

distinguish between reinforcement and balancing loops. 4. Convert core cumulative variables into stock-and-flow models. 5. Set baseline scenarios and policy scenarios. and 6. Compare the impact of different overseas internship designs on career agency formation.

Scenario setting

In terms of scenario simulation, this study sets four scenarios (Table 4): baseline scenario, enhanced university support scenario, enhanced corporate mentor scenario, and integrated support scenario. The simulation period covers the transition stages before, during, and after the internship, and compares the changing trends of career agency, employability capital, and social capital in different scenarios. The planned model validation will include unit consistency testing, extreme-condition testing, behavioral reasonableness checks, and expert review to ensure that the model can reflect the main dynamic mechanisms of Generation Z's overseas internship career development.

Table 4: Scenario Setup and Expectations

Situation	Description	Expected observation
Baseline scenario	General overseas internship program	Career agency naturally accumulates, but growth is limited
Enhanced university support scenario	Increase pre-internship training and in-internship counseling	Reduce cross-cultural pressure and improve learning transformation
Enhanced corporate mentoring scenario	Improve the quality of supervisor feedback and task design	Improve job fulfillment and career confidence
Integrated support scenario	University support, corporate mentoring, and peer networks are simultaneously improved	Generate the strongest cumulative effect on career agency

4 Research results

4.1 Causal loop analysis results

Based on the System Dynamics (SD) modeling process, this study regards Generation Z's overseas internship as a dynamic system composed of "pre-internship preparation, internship interaction, and post-internship transfer". The proposed model suggests that career agency (Akkermans & Tims, 2017; Petruzziello et al., 2022; Pham et al., 2023) is not a single psychological trait, but a dynamic ability that gradually develops through feedback, accumulation, and delay effects between university support, corporate mentoring, peer networks, cross-cultural adaptation, and career self-management behaviors.

In terms of causal loop analysis (Figure 2), this study summarizes three reinforcing loops and two balancing loops. R1 is the “Career Self-Management–Career Agency” reinforcing loop: when students actively set career goals, search for information, and seek feedback, it will enhance their career confidence and self-efficacy, further promoting subsequent career actions. Akkermans and Tims (2017) suggested that career competence can be linked to subjective career success through job crafting behavior, supporting this study to consider active career action as a career agency accumulation mechanism. R2 is the "social capital-employability capital" reinforcing loop: corporate mentoring, peers, and overseas workplace networks can provide job information, emotional support, and employment opportunities, enabling students to accumulate convertible employability capital. R3 is the “Internship Learning–Career Confidence” reinforcing loop: challenging work tasks with appropriate feedback will improve problem-solving skills and career confidence, thereby prompting students to accept higher-level tasks. Jackson and Dean (2023) suggested that different types of WIL can enhance graduates' employability, but their effectiveness is closely related to participation quality and program design, supporting the emphasis on differences in internship quality in this study.

Secondly, the model also presents two balancing loops. B1 is the "cross-cultural pressure-learning inhibition" balancing loop: language barriers, life adaptation, and workplace cultural differences can increase psychological load, reduce active interaction and learning investment, and slow down the growth of career agency. A literature review by Vu et al. (2022) suggested that international students often face cultural, institutional, and workplace integration challenges in WIL, which is consistent with the stress inhibition mechanism of this model. B2 is the "Support Resources-Stress Regulation" balancing loop: university counseling, corporate mentoring, and peer support can reduce cross-cultural pressure and maintain students' internship participation and learning momentum.

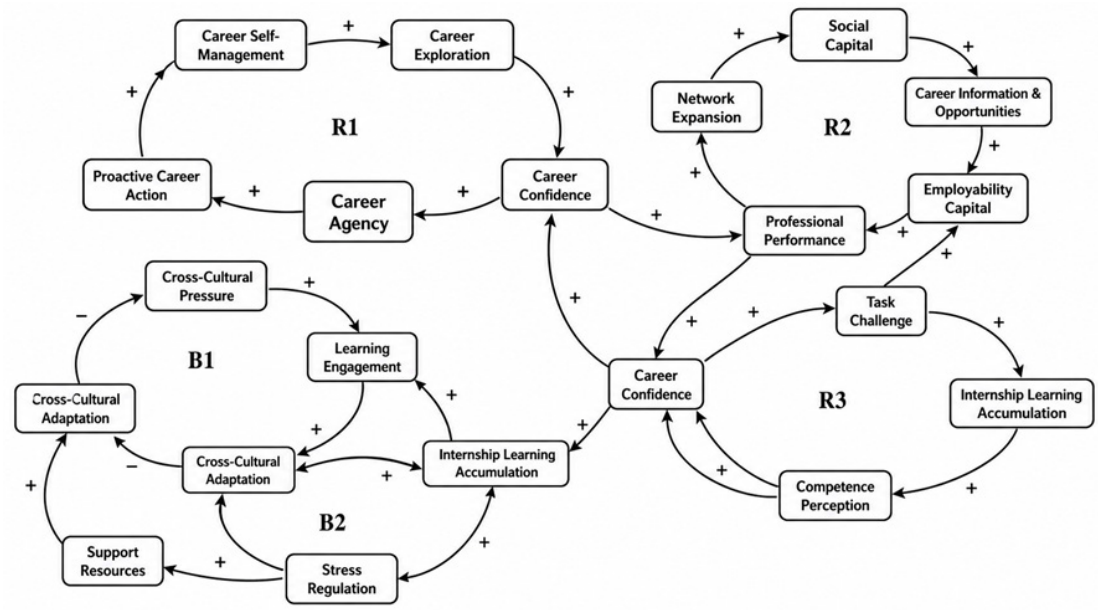


Figure 2.Causal Loop Diagram of Gen Z Overseas Internship and Career Agency Formation

4.2 Stock-and-flow model

In terms of the stock-and-flow model (Figure 3), this study sets three core stocks: career agency stock, employability capital stock, and social capital stock. The inflows into career agency include career self-management, career reflection, and proactive action experience. Outflows include career confusion, frustration, and low-quality feedback. The inflow of employability capital includes professional skills, cross-cultural competencies and internship experience; The inflow of social capital comes from mentor interaction, peer connections, and workplace networks. Luna-Reyes and Andersen (2003) suggested that qualitative data can support SD model construction and causal structure identification, so this study can further calibrate the above variable relationships through interviews and internship reflection records.

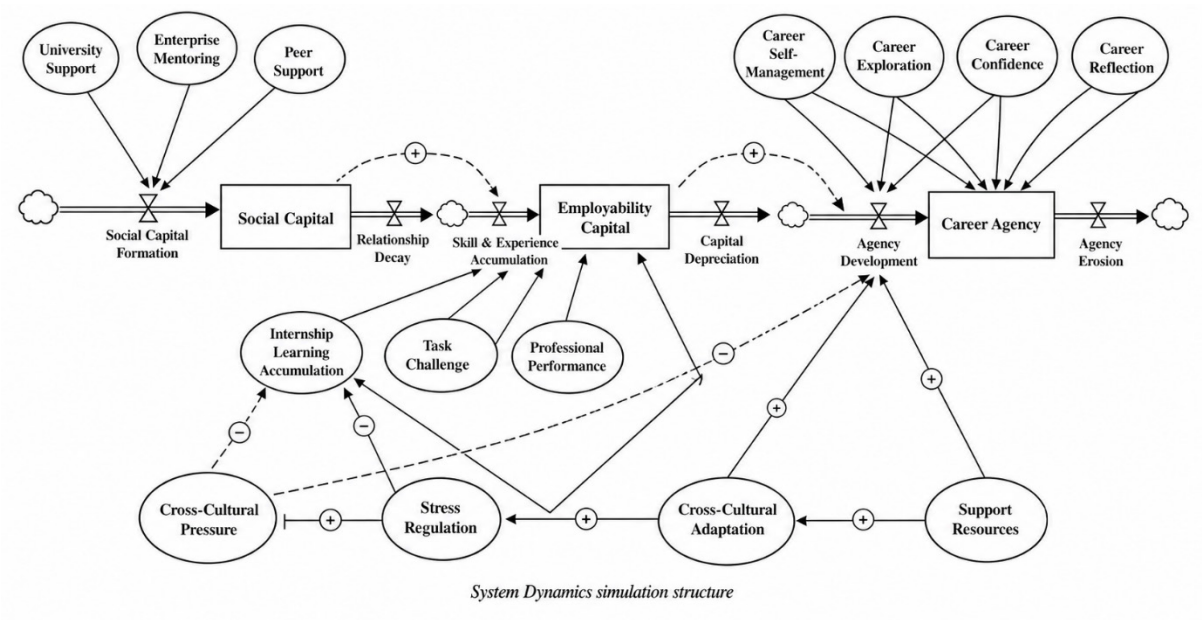


Figure 3. Stock-and-Flow Model of Career Agency Formation

4.3 Scenario simulation results

Based on the exploratory model structure, the scenario simulation (Figure 4) illustrates that under the baseline scenario, students' career agency shows a slow accumulation; Strengthening university support scenarios can reduce pre-internship uncertainty and cross-cultural pressure; strengthening the corporate

mentor scenario can improve task feedback and career confidence; In the context of integrated support, university support, corporate mentoring, and peer networks form complementary feedback, resulting in simultaneous growth of career agency, social capital, and employability capital. In terms of management implications, university overseas internship programs should not only focus on the number of dispatches, but should establish a continuous support system for pre-internship training, mentor feedback during internships, and career transformation after internships, so that overseas internships can be transformed from short-term international experience into long-term career agency development mechanisms. The scenario values are treated as illustrative assumptions and require future empirical calibration.

Table 4: Scenario simulation results

Situation	Career agency changes	Main explanation
Baseline scenario	Rise slowly	Students can naturally gain experience through internships, but the results are unstable when support is insufficient
Strengthen university support	Moderate rise	Pre-internship preparation reduces uncertainty and improves cross-cultural adaptation
Strengthen corporate mentoring	Significantly increased	High-quality tasks and feedback enhance career confidence and a sense of competence
Integrate support scenarios	Rise quickly and steadily	Universities, enterprises, and peers support each other to form a complementary feedback mechanism

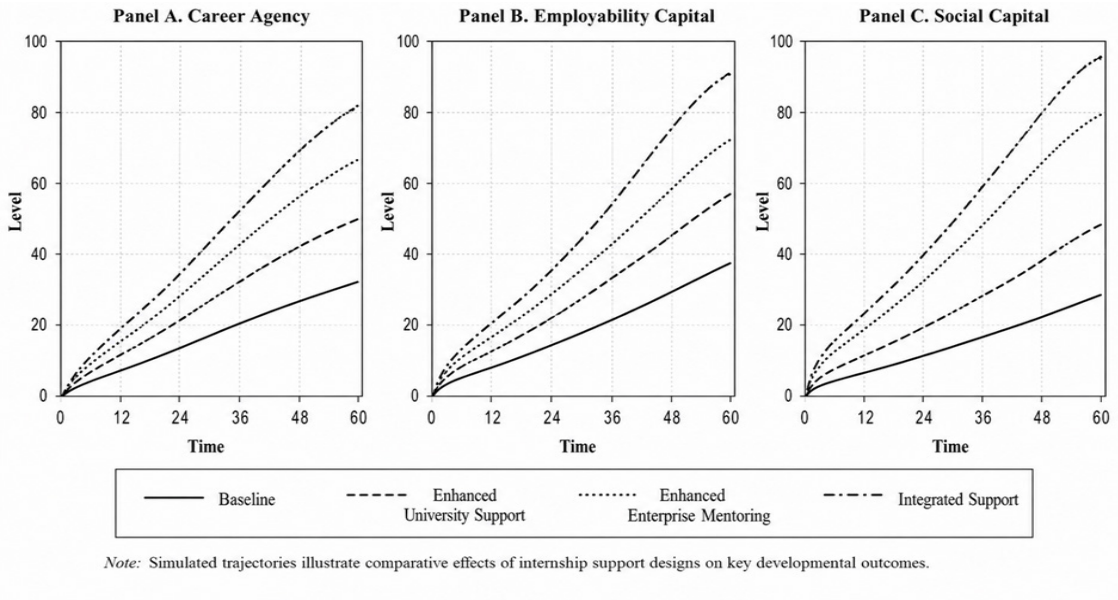


Figure 4. Scenario Simulation Results of Internship Support Designs

5 Discussion

5.1 Theoretical implications

The theoretical implications of this study are to transform the "career agency" in Generation Z's overseas internships from static psychological characteristics to dynamic abilities influenced by relationship context. Existing studies have suggested that the WIL experience of Gen Z affects their job value and career readiness (Drewery et al., 2023), and the employability agency of international graduates also needs to be formed through the combination of personal capital, social relationships, and institutional conditions (Pham et al., 2023). Therefore, this study uses System Dynamics (SD) to present the feedback relationship between career self-management, social capital, employability capital, and cross-cultural stress, which can make up for the limitations of traditional WIL research that focuses on one-way effect evaluation.

From a knowledge-management perspective, overseas internships can be viewed as structured learning systems through which experiential knowledge, relational knowledge, and career knowledge are accumulated and transformed into career agency.

5.2 Practical implications

In terms of management practice, the model-based analysis suggests that the key to the effectiveness of overseas internship programs is not the number of dispatches, but whether a continuous governance mechanism of "pre-internship preparation, internship support, and post-internship transfer" is formed. Jackson and Dean (2023) suggested that different WIL patterns may promote employability, but their effectiveness depends on program design and learning quality. Vu et al. (2022) also suggested that international students often face cultural and institutional barriers in WIL. Therefore, universities should establish cross-cultural training, corporate mentoring systems, reflective learning records, and career counseling after returning to the home country, so that overseas internships can be transformed from short-term international experience into long-term career development capital. This implication is consistent with studies on international mobility, employability agency, and Gen Z WIL experiences (Drewery et al., 2023; Pham et al., 2025).

5.3 Management and policy implications

Although ESG is not modeled as a separate stock or flow in this study, overseas internship governance reflects the social and governance dimensions of higher education sustainability and can be viewed as a concrete field in which higher education institutions fulfil their social and governance responsibilities. The introduction of ESG into higher education has been seen as an important direction for promoting sustainability paradigms and Sustainable Development Goals (SDGs) practices (Finatto et al., 2024). Therefore, overseas internship governance should emphasize equitable participation, student safety, career support, partner-institution review, learning-outcome tracking, risk management, and information transparency.

6 Conclusion

This study explores the formation mechanism of career agency in the context of overseas internship relationships of Generation Z from the perspective of System Dynamics (SD). The model-based analysis suggests that the value of overseas internships lies not only in improving short-term employability, but also in promoting the dynamic accumulation of career self-management, social capital, and employability capital through university support, corporate mentoring, peer networks, and cross-cultural adaptation. This conclusion echoes the emphasis placed on graduate employability in WIL research and complements its lack of handling feedback, delay, and nonlinear mechanisms (Jackson & Dean, 2023; Vu et al., 2022).

This study offers three preliminary contributions: first, to transform career agency from static psychological characteristics to dynamic abilities that can be accumulated through internship experience and relationship context; second, establish a causal loop and stock-and-flow structure of "career self-management-social capital-employability capital" in overseas internships; Third, it proposes a scenario simulation perspective to assist higher education institutions in comparing the long-term effects of different internship support designs. This design extends the research direction of applying CLDs and SFDs to examine the formation of career agency. Career competence and active job crafting affect individual career success, supporting this study to consider student initiative as an important mechanism for career agency growth (Akkermans & Tims, 2017).

Future research should collect longitudinal data from overseas internship students to calibrate the trajectories of career agency and employability capital, compare system structures across countries, industries, and internship types, and develop simulation-based decision-support tools for university career centers and international offices.

6 Declaration of Ethics

This study is currently based on literature-based model construction and does not involve personally identifiable human-subject data. If future surveys, interviews, or internship reflection records are collected, informed consent, confidentiality protection, and institutional ethical review will be implemented.

7 Artificial intelligence declares

This study uses ChatGPT to assist in verifying citations and optimizing arguments.

References

Akkermans, J., & Tims, M. (2017). Crafting your career: How career competencies relate to career success via job crafting. Applied Psychology, 66(1), 168–195. https://doi.org/10.1111/apps.12082

Donald, W. E., Baruch, Y., & Ashleigh, M. (2019). The undergraduate self-perception of employability: Human capital, careers advice, and career ownership. *Studies in Higher Education*, *44(4)*, 599–614.

Drewery, D., Truong, M., & Fannon, A. M. (2023). Gen Z students' work-integrated learning experiences and work values. *Higher Education, Skills and Work-Based Learning*, *13(5)*, 1023–1036.

Finatto, C. P., Fuchs, P. G., Aguiar Dutra, A. R., & Guerra, J. B. S. D. A. (2024). Environmental, social, governance and sustainable development goals: promoting sustainability in universities. *International Journal of Sustainability in Higher Education*, *25(6)*, 1121–1136.

Forrester, J. W. (1994). System dynamics, systems thinking, and soft OR. *System Dynamics Review*, *10(2–3)*, 245–256.

Jackson, D., & Dean, B. A. (2023). The contribution of different types of work-integrated learning to graduate employability. *Higher Education Research & Development*, *42(1)*, 93–110.

Lasrado, F., Dean, B., & Eady, M. (2024). University-workplace partnerships in work-integrated learning: A scoping review. *International Journal of Work-Integrated Learning*, *25(4)*, 565–602.

Luna-Reyes, L. F., & Andersen, D. L. (2003). Collecting and analyzing qualitative data for system dynamics: methods and models. *System Dynamics Review: The Journal of the System Dynamics Society*, *19(4)*, 271–296.

Patton, W., & McMahon, M. (2006). The systems theory framework of career development and counseling: Connecting theory and practice. *International Journal for the Advancement of Counselling*, *28(2)*, 153–166.

Pavlov, O. V., Cavana, R. Y., Wheat, I. D., Saeed, K., Radzicki, M. J., & Dangerfield, B. C. (2025). Introducing feedback thinking and system dynamics modeling in economics education. *System Dynamics Review*, *41(2)*, e70001.

Petruzziello, G., Antonio, A. A., Chiesa, R., & Mariani, M. G. (2022). It takes more than agency: Linking support from teaching staff, career engagement, and movement capital among university students. *Frontiers in Psychology*, *13*, 1083698.

Pham, T., Dai, K., & Saito, E. (2025). Forms of agency enacted by international Ph.D. holders in Australia and Ph.D. returnees in China to negotiate employability. *Compare: A Journal of Comparative and International Education*, *55(2)*, 192–209.

Pham, T., Jackson, D., & McIlveen, P. (2023). *Conceptualizing the employability agency of international graduates*. In Student agency and self-formation in higher education (pp. 143–165). Cham: Springer Nature Switzerland.

Richardson, G. P. (2020). Core of system dynamics. In *System dynamics: Theory and applications* (pp. 11–20). Springer.

Schroth, H. (2019). Are you ready for gen Z in the workplace?. *California Management Review*, *61(3)*, 5–18.

Sterman, J. D. (2000). *Business dynamics: Systems thinking and modeling for a complex world*. Irwin/McGraw-Hill.

Sterman, J. D. (2002). All models are wrong: reflections on becoming a systems scientist. *System Dynamics Review*, *18(4)*, 501–531.

Vu, T., Ferns, S., & Ananthram, S. (2022). Challenges to international students in work-integrated learning: A scoping review. *Higher Education Research & Development*, *41(7)*, 2473–2489.

Wilhelm, F., Hirschi, A., & Schläpfer, D. (2024). The multidimensional nature of career self-management behaviours and their relation to facets of employability. *Journal of Occupational and Organizational Psychology*, *97(1)*, 342–375.

AI-Empowers Smart Aquaculture: A System Dynamics-Integrated Decision-Making Architecture

Ju-We Chen, Shin-Ping Su
Department of Business Administration, Shih Chien University. Taipei, Taiwan

Abstract

The global aquaculture industry is facing multiple challenges, including climate change, water quality risks, energy cost fluctuations, feed efficiency pressures, and increasing demands for sustainable governance. Although technologies such as artificial intelligence of things (AIoT), smart sensors, precision feeding, automated aeration, and intelligent sludge removal have been increasingly introduced into aquaculture farms, most applications remain focused on single-point monitoring, real-time warning, or device-level optimization. Limited attention has been paid to the long-term feedback relationships among water quality, feeding strategies, fish growth, energy consumption, sludge accumulation, and environmental, social, and governance (ESG) performance. Based on this, this study constructs the "AI-empowered smart aquaculture system dynamics integrated decision-making framework", which uses system dynamics as the method basis to integrate AIoT sensing data, artificial intelligence (AI) analysis signals, smart equipment information, and aquaculture management knowledge, and establishes a causal loop diagram and stock-and-flow model. The study identified four core feedback structures, including two reinforcing loops: “AI data learning–decision quality” and “precision feeding–fish growth”, as well as two balancing loops of "overfeeding - deterioration of water quality" and "aeration improvement - energy load". Further through the comparison of baseline scenarios, AIoT monitoring scenarios, and system dynamics (SD) integrated AI decision-making scenarios, this study points out that system dynamics can help managers understand time delays, nonlinear relationships, and cumulative effects, and evaluate the impact of different feeding, aeration, sludge removal, and energy control strategies on production efficiency, environmental impact, and governance transparency. The research contribution is to upgrade smart aquaculture from technology introduction to a dynamic management framework that integrates data, knowledge, decision-making, and ESG governance, and to provide a practical reference for aquaculture operators to shift from experience-oriented to model-assisted decision-making.

Keywords: AIoT, Smart Aquaculture, System Dynamics, Decision Support, ESG Governance

Paper Type: Academic Research Paper

1 Introduction

Global aquaculture is facing multiple pressures such as climate change, energy cost fluctuations, water quality risks, feed efficiency, and sustainable governance. Although aquaculture is regarded as an important industry in response to global protein demand, its production process is highly dependent on dynamic variables such as dissolved oxygen, temperature, pH, conductivity, feeding strategy, aeration energy, and waste management. Føre et al. (2018) proposed the concept of "precision fish farming", emphasizing the need to improve aquaculture operators' monitoring, control, and recording capabilities for biological production processes through sensing, control engineering, and real-time monitoring, indicating that intelligence has become an important direction for aquaculture transformation.

However, existing smart aquaculture mostly focuses on single-point equipment or real-time monitoring, such as artificial intelligence of things (AIoT) sensors, water quality monitoring, smart feeding, automatic aeration, and sludge removal equipment. Although AIoT can support real-time environmental monitoring, data-driven decision-making, and automated management, and improve operational efficiency and production stability, most applications still focus on short-term warning or equipment-level optimization, and less deal with long-term feedback, time delay, and nonlinear relationships in aquaculture systems (Huang & Khabusi, 2025). In particular, there is an interaction between feeding rate, dissolved oxygen, fish growth, sludge accumulation, energy consumption, and environmental performance, and without an integrated decision-making model, aquaculture management may still rely on empirical judgment and is difficult to translate into a traceable, simulation-ready, and governable knowledge system.

Therefore, the problem of this study is that AI-enabled smart aquaculture should not be regarded as just the introduction of sensing technology or automation equipment, but should be understood as a dynamic management framework that integrates data, knowledge, decision-making, and governance. System dynamics is suitable for analyzing complex systems with feedback, delays, and cumulative effects. Relevant studies have applied dynamic models to recirculating aquaculture systems to evaluate the impact of water quality parameters and management strategies on fish performance and system sustainability (Udayakumar et al., 2025; Kamali et

al., 2022). At the same time, the sustainable transformation of aquaculture also needs to move beyond fragmented technical issues and move towards a systematic governance approach that can address the interaction between environmental, social, and governance (Partelow et al., 2023).

Based on the above background, this study aims to construct an "AI-empowered smart aquaculture system dynamics integrated decision-making framework", integrating AIoT sensing data, AI analysis, smart equipment information, and aquaculture management knowledge into the system dynamics model to present the causal relationship between water quality stability, precision feeding, energy use, fish growth, sludge accumulation management, and environmental, social, and governance (ESG) performance. The expected contributions of this study include: first, theoretically integrating the perspectives of smart aquaculture, AI knowledge governance, and system dynamics; Second, it proposes a decision-making framework that can conduct scenario simulation and sensitivity analysis. Third, in practice, it assists aquaculture operators from experience-oriented management to data- and model-assisted decision-making, thereby improving production efficiency, environmental performance, and governance transparency.

2 Literature review

2.1 AIoT and smart aquaculture

Smart aquaculture refers to the real-time grasp of water quality, biological growth, feeding, energy, and environmental risks through sensors, the Internet of Things, artificial intelligence, automated control, and data analysis technologies. Føre et al. (2018) proposed the concept of precision fish farming (PFF), advocating the integration of control engineering, real-time sensing, and data recording into fish production systems to enhance aquaculture operators' ability to monitor, control, and record biological production processes, and reduce the dependence of traditional aquaculture on human experience and subjective judgment. This view lays an important foundation for smart aquaculture to move from "equipment automation" to "data-driven decision-making".

In recent years, the application of AIoT in aquaculture has expanded rapidly, including water quality monitoring, fish behavior identification, disease detection, precision feeding, automatic aeration, and predictive maintenance. Huang and Khabusi (2025) pointed out that AIoT combines Internet of Things (IoT), and AI algorithms to collect real-time data such as temperature, pH, dissolved oxygen, salinity, and fish behavior, and supports water quality management, disease early warning, and automated control through predictive analytics. However, existing research mostly focuses on sensing accuracy, algorithm performance, or individual equipment efficiency, and rarely further explores how multi-source data can be translated into medium- and long-term business decisions, governance capabilities, and sustainable performance of farms. Therefore, the core value of AIoT in smart aquaculture lies not only in real-time monitoring but also in the formation of knowledge assets that can be accumulated, explained, and governed.

2.2 Systemic complexity of intelligent farming decisions

Aquaculture is not a simple production process, but a complex adaptive system composed of water quality, biological growth, feed input, energy use, waste accumulation, and environmental conditions. Aquaculture managers must weigh production efficiency, fish health, energy costs, and environmental impact. For example, increasing feeding may promote fish growth in the short term, but overfeeding can also lead to increased bait residue, sludge accumulation, deterioration of water quality, and increased mortality. Increasing aeration intensity can improve dissolved oxygen, but at the same time increase energy use and carbon emissions. Kamali et al. (2022) established a dynamic model for recirculating aquaculture systems, pointing out that water quality parameters and management strategies jointly affect fish performance, and aquaculture systems need to understand the relationship between management behavior and production results through dynamic modeling.

This feature shows that smart aquaculture decisions cannot rely solely on single-point monitoring or short-term warnings, but need to deal with time delays, nonlinear relationships, and cumulative effects. For example, the impact of dissolved oxygen decline on fish health may be delayed, and the effect of sludge accumulation on water quality deterioration may not be apparent until a period of time. Without an integrated model, AIoT devices can generate a large amount of data, but it is still difficult for managers to judge the medium- and long-term effects of different decision-making solutions. Therefore, smart aquaculture needs to be further developed from "sensing-oriented" to "systematic decision-oriented".

2.3 Theoretical basis of system dynamics

System dynamics (SD) was initially developed by Forrester (1961) to analyze complex industrial and management systems with feedback structures, time delays, stock accumulation, and nonlinear behavior. Sterman (2000) pointed out that the root cause of many management problems is not a single event, but a dynamic behavior formed by the system's internal feedback structure and decision-making delays. Therefore, SD is particularly suitable for analyzing systems such as aquaculture that are affected by the interaction of biological, technological, environmental, and management factors, and transforming the "data-knowledge-decision-performance" in aquaculture systems into a simulation-ready dynamic architecture.

In smart aquaculture research, SD can make up for the shortcomings of AI and IoT's emphasis on real-time prediction. AI models can provide water quality predictions, anomaly detection, and operational recommendations, but SD can further integrate these signals into the overall feedback structure to analyze the long-term impact of different management strategies. Causal loop diagram (CLD): Shows the reinforcing and balancing loops between variables. Stock-and-Flow Diagram (SFD): Shows the stock and flow of water quality, fish biomass, energy accumulation consumption, sludge accumulation and governance capability. Recent studies have applied SD to water quality management in recirculating aquaculture systems, integrating the relationship between fish metabolism, feed input, biological filter, water exchange, aeration, and water quality parameters through models to evaluate the effects of different aquaculture densities, feeding intensities, and aeration strategies on system performance (Udayakumar et al., 2025). Therefore, SD is not only a simulation tool, but also a method of translating sensing data, management experience, and strategic assumptions into a communicative decision-making framework.

2.4 ESG governance and sustainable transformation of smart aquaculture

With the emphasis on sustainable development and ESG governance, the performance evaluation of aquaculture is no longer limited to production and revenue, but extends to water resource use, energy efficiency, carbon emissions, pollution control, food safety, social responsibility, and governance transparency. Partelow et al. (2023) point out that the sustainable transformation of aquaculture requires a greater emphasis on governance and a shift from fragmented technical issues to a systematic approach that addresses social, environmental, and institutional interactions. Therefore, the sustainable value of smart aquaculture lies not only in improving production efficiency but also in supporting traceable, verifiable, and governable operational decisions.

From an ESG perspective, the integration of AIoT with SD provides three governance functions (Table 1). First, in terms of the environment, sensing data can support water quality stability, energy efficiency, and pollution load management. Second, in terms of society, smart decision-making can reduce the risk of aquaculture failure, improve food safety, and narrow the knowledge gap among small and medium-sized aquaculture operators. Third, in terms of governance, the system model can transform decision-making basis, risk warning, and management results into traceable records, improving decision-making transparency and responsibility management capabilities. It can be seen that if AI-enabled smart aquaculture can be combined with SD, it can not only improve operational efficiency but also form an ESG-oriented knowledge governance framework.

Table 1: ESG Dimensions and System Dynamics Variables in Smart Aquaculture Governance

Dimension	Governance focus	System dynamics variables
Environmental	Water quality stability, pollution load, sludge accumulation, energy consumption, and carbon pressure are used as environmental performance metrics.	DO, pH, energy consumption, sludge accumulation
Social	Using fish survival, food safety risk reduction, production stability, and farmer knowledge as social value proxies	Fish health, mortality, and operational knowledge and ability
Governance	Using decision records, ESG dashboards, data traceability, and decision response time as governance transparency proxies	Data visualization, AI decision-making signals, and governance response speed

2.5 Literature gap

Based on the above literature, existing research has confirmed that AIoT, precision aquaculture, and automated equipment have important value for aquaculture, but there are still three gaps. First, existing research mostly focuses on a single technology or single device, and rarely integrates the dynamic relationship between water quality, feeding, energy, fish growth, and pollution accumulation. Second, AIoT research

emphasizes real-time monitoring and prediction capabilities, and less explores how data can be transformed into organizational knowledge, governance capabilities, and medium- and long-term decision support. Third, ESG research is more biased towards performance disclosure and policy discussions, and less often embeds ESG indicators into farm daily operations and scenario simulations. Therefore, this study uses system dynamics as an integrated approach to construct an AI-empowered smart aquaculture decision-making framework to respond to the research needs of smart aquaculture from technology introduction to knowledge governance and sustainable decision-making.

3 Research Methodology

This study uses SD as the main research method to construct an "AI-empowered smart aquaculture system dynamics integrated decision-making framework". System dynamics is suitable for analyzing complex management systems with feedback relationships, time delays, stock accumulation, and nonlinear behavior. Sterman (2000) pointed out that many organizational and management problems do not stem from a single event, but from dynamic behaviors formed by feedback structures and decision delays within the system. Therefore, this study uses SD to link AIoT data, smart devices, aquaculture decision-making, and ESG governance performance to make up for the shortcomings of traditional smart aquaculture research focusing on single-point monitoring and real-time control.

3.1 Study design

The modeling process of this study is divided into five steps. First, literature analysis: Define the boundaries of the system and set the research scope to the operation decision-making system of smart aquaculture farms, covering water quality management, feeding management, aeration energy, fish growth, sludge accumulation, smart equipment investment, and ESG governance. Second, carry out variable identification, based on the literature on precision fish farming and AIoT smart aquaculture, extract core variables such as dissolved oxygen, temperature, pH, feeding accuracy, fish biomass, mortality, energy consumption, and AI prediction accuracy. Føre et al. (2018) pointed out that the purpose of precision fish farming is to improve the ability of aquaculture operators to monitor, control, and record biological production processes through control engineering principles. This view provides a theoretical basis for the establishment of data-driven variables in this study (Føre et al., 2018). Third, construct a causal loop diagram (CLD) to identify the reinforcing and balancing loops between AI data learning, precision feeding, dissolved oxygen control, sludge accumulation, and management response. Fourth, establish a stock-and-flow diagram (SFD) that sets fish biomass, water quality stability, sludge accumulation, energy consumption, and governance capability as the main stocks, and uses growth rate, pollution load, sludge removal efficiency, energy regulation, and data learning rate as flows. Fifth, through scenario simulation and sensitivity analysis, compare the medium- and long-term effects of different intelligent management strategies.

3.2 Research Framework

In terms of variable design, this study divides systematic variables into six categories. The first is water quality variables, including dissolved oxygen, temperature, pH, conductivity and redox potential; the second is production variables, including fish biomass, growth rate, feed conversion rate and mortality; the third is intelligent equipment variables, including AI prediction accuracy, AIoT data quality, intelligent feeding efficiency, aeration control intensity, and sludge removal efficiency; the fourth is environmental variables, including sludge accumulation, pollution load, energy consumption and carbon emissions; the fifth is management variables, including decision-making response time, aquaculture operators' knowledge ability and system adoption; The sixth is ESG governance variables, including environmental performance, social value, and governance transparency. Kamali et al. (2022) established a dynamic model for recirculating aquaculture systems, pointing out that water quality parameters and management strategies jointly affect fish body performance, and can evaluate system behavior under different management strategies through dynamic simulation, supporting the core model structure of this study, which supports the core model structure of water quality, biological growth, and management behavior. The research structure is summarized in Table 2.

Table 2: Architecture Layers of the Proposed Decision-Making Framework

Levels	Contents	Function
Data layer	AIoT sensors, water quality meters, smart feeding, sludge removal equipment, energy equipment	Collect real-time operational data
Knowledge layer	AI analysis, water quality prediction, anomaly detection, rules of thumb	Transform data into decision-making knowledge
System dynamics layer	CLD, SFD, Scenario Simulation, Sensitivity Analysis	Analyze long-term feedback and policy effects
Governance layer	ESG indicators, decision records, risk warnings, and management dashboards	Support transparent governance and sustainable management

The data sources of this study include three categories. The first is academic literature, mainly taken from research related to smart aquaculture, AIoT, precision aquaculture, system dynamics, and ESG governance, and is used to establish variable relationships and model assumptions. The second is the data of smart aquaculture equipment, including the operation data generated by AIoT sensors, water quality monitoring equipment, intelligent feeding systems, aeration equipment, and sludge removal systems, such as dissolved oxygen, temperature, pH, feeding volume, equipment operating time, and energy consumption. The third is management and governance data, including farmer operation records, decision response time, abnormal event handling records, energy costs, and ESG performance indicators. Huang and Khabusi (2025) pointed out that AIoT can integrate sensors, AI algorithms, and automation equipment in aquaculture for water quality management, disease early warning, and decision support, indicating that multi-source data integration is an important foundation for smart aquaculture decision-making architecture. The core variables are summarized in Table 3.

Table 3: Core Variables of the Smart Aquaculture System Dynamics Model

category	Core variables	Variable description	SD role
Water quality management	Dissolved Oxygen, DO	The concentration of dissolved oxygen in the aquaculture water body for fish to breathe is a key water quality indicator that affects the health and mortality of fish.	Auxiliary variable/ state variables
Water quality management	Temperature	Water temperature affects fish metabolism, feeding behavior, dissolved oxygen requirements, and disease risk.	Auxiliary variable
Water quality management	pH	It reflects the pH state of the water body, affecting the physiological adaptation of fish, microbial action and ammonia nitrogen toxicity.	Auxiliary variable
Water quality management	Electrical Conductivity, EC	It reflects changes in water ion concentration and salinity, and can be used as a reference index for water quality changes and pollution loads.	Auxiliary variable
Water quality management	Oxidation-Reduction Potential, ORP	It reflects the redox state of water bodies, which is related to the decomposition of organic matter, microbial activities, and water quality safety.	Auxiliary variable
Production performance	Fish Biomass	The total biomass of fish in the pond is affected by growth, death, and harvest.	Stock
Production performance	Fish Growth Rate	The growth rate of fish is affected by feed efficiency, water quality stability, and fish health.	Flow
Production performance	Mortality Rate	Fish mortality is affected by low dissolved oxygen, deterioration of water quality, disease and delayed management.	Flow
Production performance	Feed Conversion Efficiency	Feed conversion efficiency refers to the efficiency of the input feed into fish growth.	Auxiliary variable
Feeding management	Feeding Amount	The amount of feed input per unit time.	Flow / Decision variable
Feeding management	Feeding Precision	The feeding system accurately feeds the fish according to the needs, behavior, and water quality conditions.	decision variable
Pollution and sludge accumulation	Feed Residue and Waste	The organic load formed by uningested feed and fish excrement.	Flow

category	Core variables	Variable description	SD role
Pollution and sludge accumulation	Sludge Accumulation	The accumulation of organic matter, residual bait and excrement at the bottom of the aquaculture pond.	Stock
Pollution and sludge accumulation	Smart Cleaning Removal	The efficiency of smart sludge removal or automatic cleaning equipment to remove sludge accumulation and contaminants.	Flow
Energy management	Aeration Intensity	The operating strength of the aeration equipment is used to improve the circulation of dissolved oxygen and water.	decision variable
Energy management	Energy Consumption	The total energy use of aeration, water circulation, monitoring, and automation equipment.	Stock
Energy management	Energy Control	Adjust the control capability of equipment operation based on water quality, dissolved oxygen demand, and energy costs.	decision variable
AI and data	AIoT Data Volume	The amount of aquaculture operation data accumulated by sensors, equipment, and management systems.	Stock / Auxiliary variable
AI and data	AI Prediction Accuracy	The accuracy of AI models in predicting water quality changes, fish growth, feeding needs, and abnormal events.	Auxiliary variable
AI and data	Anomaly Detection Capability	The ability of the system to identify abnormal water quality, equipment or fish behavior.	Auxiliary variable
Management decisions	Decision Quality	The ability of farm managers or systems to make appropriate decisions based on data and models.	Auxiliary variable
Management decisions	Decision Response Time	The time it takes to manage the execution of an action from the occurrence of an anomaly.	Auxiliary variable
Management decisions	Farmer Knowledge	Aquaculture operators' understanding and application of water quality, biological growth, equipment operation, and AI suggestions.	Stock / Soft variable
Governance and sustainability	Governance Capability	The farm's ability to institutionalize and manage data, decision-making records, risk warnings, and ESG indicators.	Stock
Governance and sustainability	ESG Dashboard	A management interface that visualizes and tracks environmental, social, and governance metrics.	Auxiliary variable
Governance and sustainability	Environmental Performance	environmental performance such as water quality stability, pollution reduction, energy efficiency and carbon emission control.	Outcome variable
Governance and sustainability	Social Value	social values such as food safety, stable production, reduced aquaculture risks, and narrowing the knowledge gap of small and medium-sized businesses.	Outcome variable
Governance and sustainability	Governance Transparency	decision-making records, data traceability, risk disclosure, and ESG reporting capabilities.	Outcome variable

Scenario simulation design is divided into three scenarios. The first is the baseline scenario, which represents traditional experiential management, which mainly relies on manual observation and aquaculture experience for feeding, aeration, and water quality treatment. The second is the AIoT monitoring scenario, which represents the introduction of sensors and smart devices, but the decision-making is still based on manual judgment. The third is the SD-integrated AI decision scenario, which represent the integration of AI prediction signals, smart device data, and system dynamics model to support feeding, aeration, sludge removal, and energy management decisions through scenario simulation. The three scenarios will compare indicators such as water quality stability, fish growth rate, mortality, feed efficiency, energy consumption, pollution load, and ESG governance transparency. Partelow et al. (2023) pointed out that the sustainable transformation of aquaculture requires a shift from fragmented technical issues to a systematic governance approach that can address interconnected sustainability issues.

Overall, this research method does not only pursue short-term prediction accuracy with AI models, but also transforms AIoT data into interpretable, simulation-ready, and governable decision-making knowledge through SD. This approach helps uncover key leverage points in smart aquaculture systems and assists aquaculture operators in assessing the long-term impact of different management strategies on production efficiency, environmental impact, and ESG governance capabilities.

4 Research results

4.1 AI empowers smart aquaculture to integrate decision-making framework

The proposed architecture integrates five layers: data acquisition, AI analytics and knowledge, system dynamics modeling, decision and governance, and operational and sustainability outcomes (Figure 1) to illustrate the dynamic relationship between AIoT data, smart devices, aquaculture management decisions, and ESG performance. The qualitative comparison suggests that the key to smart aquaculture lies not only in the introduction of sensors, AI algorithms, or automated equipment, but also in the ability to transform disparate data into interpretable, simulation-ready, and governable management knowledge. This view echoes the concept of precision fish farming proposed by Føre et al. (2018), that is, to improve the monitoring, control, and recording capabilities of aquaculture operators in biological production processes through sensing, monitoring, and control engineering.

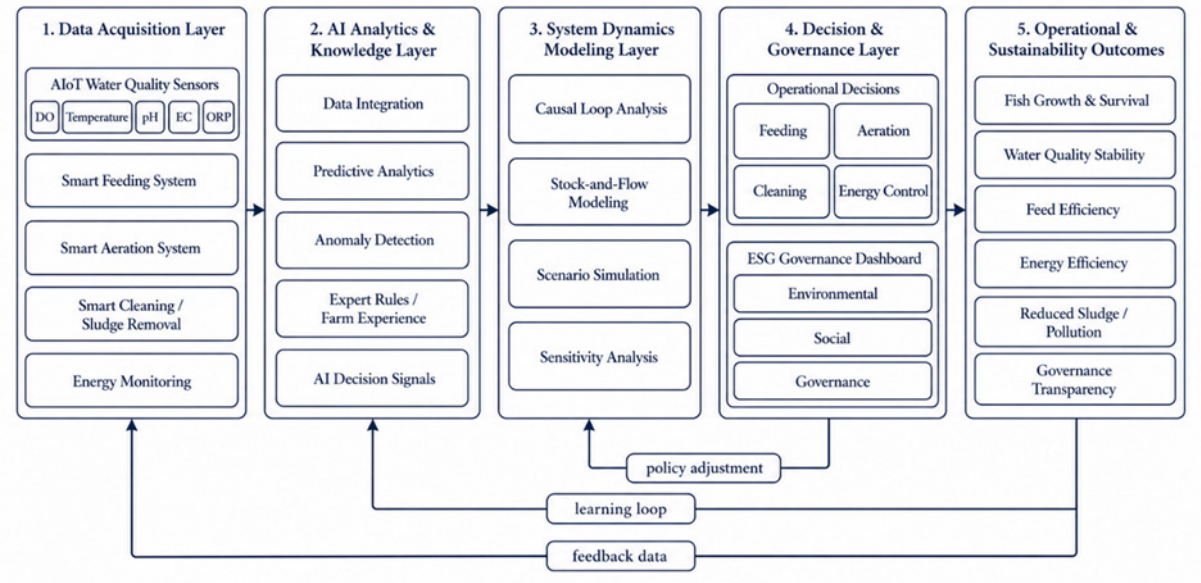


Figure 1. AI-Empowered Smart Aquaculture Decision-Making Architecture

4.2 Causal loop analysis results

The causal loop analysis identifies two reinforcing loops and two balancing loops. R1 explains how AIoT data accumulation improves AI prediction accuracy and decision quality, which further enhances aquaculture performance and technology adoption. R2 captures how precision feeding improves feed conversion efficiency, fish growth, farm revenue, and reinvestment in smart technologies. In contrast, B1 shows that overfeeding increases feed residue, sludge accumulation, water quality deterioration, and mortality risk, thereby triggering feeding adjustment. B2 indicates that aeration improves dissolved oxygen and fish survival but also increases energy consumption and carbon pressure, requiring energy-control adjustment(Figure 2).

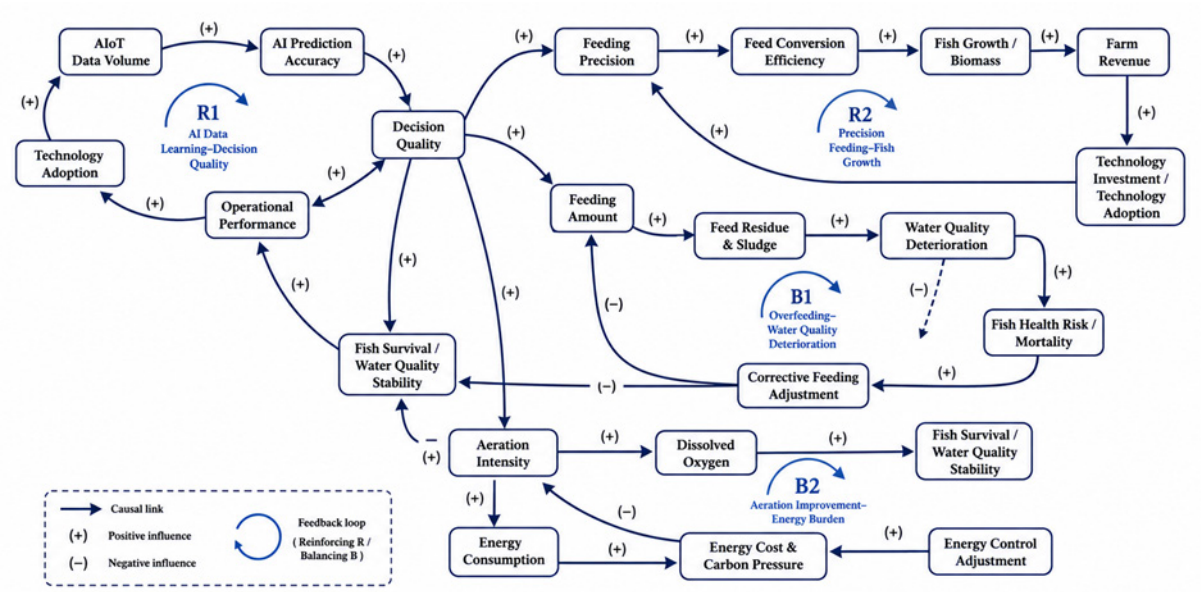


Figure 2: Causal Loop Diagram of AI-Empowered Smart Aquaculture System

4.3 Stock-and-flow model results

In terms of Stock-and-flow model, the smart aquaculture system is designed into five main stocks: fish biomass, water quality stability, sludge accumulation, energy consumption, and governance capability. Fish biomass is affected by growth rate, mortality and harvest rate; Water quality stability is affected by dissolved oxygen, pH, temperature, pollution load, and water treatment efficiency. Sludge accumulation is determined by feed residue, fish waste, and sludge removal efficiency. Energy consumption mainly comes from the operation of aeration, circulating water equipment and smart equipment; Governance capabilities are accumulated from data quality, decision-making records, AI prediction accuracy, and ESG disclosure capabilities. The dynamic model of recirculating aquaculture by Kamali et al. (2022) shows that management strategies and water quality parameters can jointly affect fish performance, and the system behavior under different strategies can be evaluated through dynamic simulation. This is consistent with the design of this study to integrate production, water quality, and management behaviors into the SFD framework.

4.4 Scenario simulation results

This study designs three qualitative scenarios to compare different smart aquaculture management logics: the baseline scenario, the AIoT monitoring scenario, and the SD-integrated AI decision scenario(Table 4). Since this research proposes a conceptual decision-making architecture rather than a calibrated numerical simulation model, the scenario comparison is presented qualitatively to illustrate how data integration and system modelling may influence production, environmental management, and ESG governance.

In the baseline scenario, aquaculture operations depend mainly on manual observation and experience-based judgment. Although flexible, this approach is constrained by delayed responses, incomplete information, and fragmented records, resulting in lower water quality stability, higher fish mortality, weaker feed and energy efficiency, greater sludge and pollution loads, and lower ESG transparency. The AIoT monitoring scenario improves real-time data collection, water quality monitoring, abnormality detection, and decision response speed through sensors and smart devices. However, without system dynamics modelling, it remains limited to short-term monitoring and warning, making it difficult to evaluate medium- and long-term effects.

The SD-integrated AI decision scenario shows the strongest expected performance. By combining AI prediction signals, AIoT sensing data, smart equipment information, and system dynamics simulation, it supports more comprehensive decisions on feeding, aeration, sludge removal, and energy control. This scenario is expected to improve water quality stability, reduce mortality, enhance feed efficiency, strengthen decision quality, and increase ESG transparency. Overall, the comparison suggests that smart aquaculture should move beyond equipment adoption toward integrated data, model-assisted decision-making, and long-term governance capability.

Table 4: Expected Scenario Comparison Results

Indicators	Baseline	AIoT Monitoring	SD-Integrated AI Decision
Water quality stability	low	Medium	High
Fish mortality	High	Medium	Low
Feed use efficiency	low	Medium	High
Energy efficiency	low	Medium	Medium-high
decision response speed	Slow	Medium	Fast
ESG transparency	low	Medium	High
LT* Governance capabilities	low	Medium	High

LT* : Long-term

4.5 Key Sensitivity Variables

Based on the qualitative scenario comparison, this study identifies three variables expected to exert strong leverage on system behavior: AI prediction accuracy, feeding precision, and aeration energy efficiency. These variables are critical because they influence decision response quality, feed conversion efficiency, water quality stability, energy costs, and carbon pressure, as summarized in Table 5.

Table 5: Key Sensitivity Variables and Expected Impacts

Key Sensitivity Variables	Impact
AI Prediction Accuracy	Improve water quality early warning, feeding decision-making and governance response
Feeding Precision	Improve feed efficiency, fish growth and reduce pollution load
Aeration Energy Efficiency	Improve dissolved oxygen stability and reduce energy costs and carbon emission pressure.

5 Discussion

This study highlights that the value of AI-empowered smart aquaculture lies in transforming fragmented operational data into decision-relevant knowledge. While AIoT sensors and smart equipment improve real-time monitoring, their managerial value remains limited if data are not connected to feedback structures, time delays, and cumulative system effects. By applying system dynamics, this study positions smart aquaculture as a knowledge-based decision system rather than a collection of isolated technologies.

From a theoretical perspective, the proposed framework extends the precision fish farming perspective by linking sensing, monitoring, and control functions with system-level learning. It shows that AI-generated data can support not only operational control but also feedback analysis, scenario learning, and strategic decision-making. This strengthens the connection between smart aquaculture, knowledge assets, and system dynamics.

From a managerial perspective, the framework suggests that aquaculture operators should move beyond experience-based responses and adopt model-assisted decision-making. For example, feeding, aeration, sludge removal, and energy-control decisions should be evaluated not only by their immediate operational effects but also by their delayed impacts on water quality, fish growth, mortality risk, energy cost, and pollution load. This helps reduce local optimization and supports more balanced operational decisions.

From an ESG governance perspective, the framework embeds environmental, social, and governance concerns into daily aquaculture operations. Water quality stability, energy consumption, pollution load, fish health, decision records, and data traceability are treated as decision variables rather than external reporting items. Therefore, AI and system dynamics integration can support sustainable governance, risk disclosure, and transparent knowledge management in smart aquaculture.

6 Conclusion

This study proposes an AI-Empowered Smart Aquaculture System Dynamics Integrated Decision-Making Framework to explain how AIoT sensing data, AI prediction signals, smart equipment information, and aquaculture management knowledge can be integrated into a system dynamics-based decision architecture. The main contribution is to reframe smart aquaculture from a technology-adoption issue into a knowledge-governance and decision-support issue. Rather than treating AIoT sensors, precision feeding, automated aeration, and sludge removal systems as isolated tools, this study connects them through causal loop analysis, stock-and-flow modeling, and qualitative scenario comparison to support medium- and long-term decision-making.

Theoretically, this study integrates AIoT-enabled smart aquaculture, precision fish farming, system dynamics, and ESG governance. It argues that AI-generated operational data should be viewed not merely as monitoring outputs, but as knowledge assets that support feedback analysis, strategic simulation, and governance transparency. Methodologically, the framework identifies two reinforcing loops and two balancing loops that clarify the dynamic relationships among water quality stability, feeding precision, fish growth, sludge accumulation, energy consumption, and ESG performance. Practically, the framework helps aquaculture operators shift from experience-based management to model-assisted decision-making.

This study is limited by its conceptual nature. The model has not yet been calibrated or validated with long-term field data, and the scenario comparison remains qualitative rather than numerical. Model assumptions may also vary across fish species, farming densities, production systems, equipment specifications, climate conditions, and operators' capabilities, which may affect generalizability.

Future research should incorporate longitudinal AIoT field data to estimate parameters, calibrate equations, and validate simulation results. Further studies may compare pond aquaculture, recirculating aquaculture systems, and offshore aquaculture to examine ESG impacts under different contexts. Future work may also develop a decision dashboard linking real-time AIoT data, system dynamics simulation, and ESG indicators as a practical decision-support tool for sustainable smart aquaculture.

7 Declaration of Ethics

This study does not involve human participants, personal data, animal experiments, or sensitive information. Therefore, no specific research ethics approval was required.

8 AI Declaration

This study used ChatGPT to assist in identifying academic references relevant to the research arguments and to support the preparation of figures. All cited references were reviewed by the authors to confirm their relevance to the study. The authors remain responsible for the accuracy, interpretation, and final content of the manuscript.

References

Forrester, J. W. (1961). *Industrial dynamics*. MIT Press.

Føre, M., Frank, K., Norton, T., Svendsen, E., Alfredsen, J. A., Dempster, T., ... & Berckmans, D. (2018). Precision fish farming: A new framework to improve production in aquaculture. *Biosystems Engineering*, *173*, 176-193. https://doi.org/10.1016/j.biosystemseng.2017.10.014

Huang, Y. P., & Khabusi, S. P. (2025). Artificial intelligence of things (AIoT) advances in aquaculture: a review. *Processes*, *13(1)*, 73. https://doi.org/10.3390/pr13010073

Kamali, S., Ward, V. C., & Ricardez-Sandoval, L. (2022). Dynamic modeling of recirculating aquaculture systems: Effect of management strategies and water quality parameters on fish performance. *Aquacultural Engineering*, *99*, 102294. https://doi.org/10.1016/j.aquaeng.2022.102294

Partelow, S., Asif, F., Béné, C., Bush, S., Manlosa, A. O., Nagel, B., ... & Turchini, G. M. (2023). Aquaculture governance: five engagement arenas for sustainability transformation. *Current Opinion in Environmental Sustainability*, *65*, 101379. https://doi.org/10.1016/j.cosust.2023.101379

Sterman, J. D. (2000). *Business dynamics: Systems thinking and modeling for a complex world*. Irwin/McGraw-Hill.

Udayakumar, R., Kadirov, I., Radjabova, D., Fallah, M. H., Tursunov, M., & Masalieva, O. (2025). A System Dynamics Model for Water Quality Management in Recirculating Aquaculture Systems (RAS). *Natural and Engineering Sciences*, *10(2)*, 434-446.

Inclusive Governance of Sport Centers: A System Dynamics Perspective

Yi-Chen Cheng, Ching-Yeh Chu
Department of Business Administration, Shih Chien University. Taipei, Taiwan

Abstract

This study focuses on "Inclusive Governance of Sports Centers" and explores the dynamic relationship between ESG governance, public services, and sustainable operations in urban public sports centers from the perspective of system dynamics. In recent years, the sports center has transformed from a mere venue and course service provider to an important social infrastructure that promotes public health, social participation, community connections, and access to the disadvantaged. However, in a service system that involves outsourcing operations, professional sports organizations, individual coaches, and diverse users, sports centers also face multiple tensions such as operational efficiency, service capacity, fair participation, governance transparency, and financial sustainability. This study adopts a conceptual system dynamics research design to construct a research framework, causal loop diagram, and stock flow model for inclusive governance in sports centers, and further compares five scenarios: baseline scenarios, inclusive curriculum design, capacity allocation, transparent adaptive governance, and cost pressure. The results show that there is a mutually reinforcing and constraint relationship between participation diversity, service capacity, service satisfaction, trust capital, governance capabilities, and resource reinvestment. Among them, transparent adaptive governance can best improve long-term trust and overall ESG performance. capacity configuration helps reduce congestion and improve service satisfaction; Inclusive curriculum design can enhance participation diversity, but if there is a lack of capacity support, it may have a service dilution effect. This research contribution is to integrate ESG, inclusive governance, and system dynamics into a decision support framework for public sports service governance, and to provide sports center managers and policymakers with a governance mindset that balances public value, operational sustainability, and social inclusion.

Keywords: Sports Center, Inclusive Governance, Systems Dynamics, ESG, Sustainable Sports Services

Paper type: Academic Research Paper

1 Introduction

In recent years, the role of urban sports centers has gradually transformed from mere sports venues and course service providers to important social infrastructure that promotes public health, social participation, community connections, and sustainable governance. Especially in the context of aging, health promotion, and increased awareness of ESG governance, sports centers not only need to meet the fitness needs of the general public but also respond to the participation needs of different users such as teenagers, the elderly, beginners, people with disabilities, and vulnerable groups. Existing studies have pointed out that sports participation systems involve multi-level interactions such as policies, facilities, managers, and users, and the governance goals of sports facilities are often affected by both public value and business performance (Kumar et al., 2018). The Taipei City Sports Center has the diverse operational characteristics of outsourcing operators, professional sports organizations, and individual coaches jointly providing services, so its governance issues cannot be solved by a single manager, but involve the coordination and resource allocation of multiple stakeholders.

While pursuing utilization, course revenue, and service diversification, sports centers may also face structural contradictions in inclusive governance. For example, popular hours and courses can cause crowded venues, making it more difficult for the elderly, beginners, or vulnerable groups to access services. Excessive pursuit of revenue may also compress public welfare, accessibility, and diversity of participation. Research on sports governance has pointed out that sports organizational governance not only involves internal organizational management but also encompasses multiple aspects such as institutional, political, social and stakeholder interactions (Dowling et al., 2018). The collaborative governance perspective also emphasizes that the public sector, civil society organizations, and community actors must share information, resources, and capabilities to handle public service tasks that are difficult for a single organization to accomplish independently (Shilbury et al., 2016). In addition, social inclusion research reminds that although sports participation itself contributes to social connections, increasing participation does not necessarily lead to true equity and inclusion without institutional design and environmental adjustment (Jeanes et al., 2021; Coalter, 2017）。

Based on the above awareness, this study believes that inclusive governance in sports centers should be understood as a dynamic system with feedback, delay, accumulation, and nonlinear characteristics, rather than

static institutional arrangements or single policy interventions. System dynamics emphasizes that system behavior comes from the interaction between feedback loops, stock traffic, and decision-making rules, making it suitable for analyzing long-term governance effects in complex public service systems (Sterman et al., 2015). Therefore, the purpose of this study is threefold: first, to construct a systematic dynamic analysis framework for inclusive governance in sports centers; second, identify the main causal loops between participation diversity, service capacity, satisfaction, trust, governance transparency, and resource reinvestment; Third, scenario simulation analyzes how different governance strategies affect the inclusive participation, service quality, and sustainable operation of sports centers. In this way, this study hopes to supplement the shortcomings of existing sports governance research that focuses on static analysis, and provide decision-making support perspectives for sports center managers and policymakers that take into account ESG, public value, and operational sustainability.

2 Literature review

2.1 Sports Center Governance and Public Sports Services

The sports center has the characteristics of both public services and market operations, and its governance involves not only venue management, curriculum design, and operational efficiency, but also public health, community participation, and social equity. Sports governance research points out that sports organizational governance has expanded from traditional internal management to include multi-level issues such as institutional arrangements, stakeholder coordination, accountability mechanisms, and public value creation (Dowling et al., 2018). In the scenario of Taiwan's urban sports center, the government, outsourced operators, professional sports organizations, and individual coaches jointly form a complex service system, shifting the focus of governance from a single operational performance to diversified service supply, resource allocation, and community responsiveness. Sports centers need to establish coordination mechanisms between operators, professional organizations, coaches, and diverse users to respond to the needs of inclusive governance.

2.2 Inclusive governance and social participation

Inclusive governance emphasizes that different social groups can access, use, and continue to participate in public services fairly. Although exercise is often seen as an important tool for promoting social connections and health and well-being, relevant research reminds us that sports participation itself does not necessarily lead to social inclusion. Without institutional design, resource support, and barrier to participation, the sports field may still reproduce gender, age, class, or ability differences (Coalter, 2017; Jeanes et al., 2021). Therefore, inclusive governance in sports centers should focus on accessibility, affordability, diversity of participation, and service responsiveness, and reduce the risk of exclusion of specific groups through transparent decision-making and user feedback (Table 1).

Table 1: Research Aspects of Inclusive Governance in Sports Centers

face	Description
Accessibility	venue, fees, time slots, transportation, and information accessibility
Participation diversity	Teenagers, the elderly, beginners, people with disabilities, general athletes, etc. participated in a variety of activities
Service responsiveness	Responsiveness to the needs of different groups
Stakeholder coordination	Coordination between governments, operators, coaches, communities, and users
Trust and transparency	Transparent decision-making, reasonable fees, and fair venue allocation

2.3 ESG and sustainable sports services

In recent years, sustainability has become an important topic in sports management research. Movement organizations not only need to deal with environmental impacts but also face challenges such as social responsibility, governance transparency, and stakeholder trust. The sustainable sports management literature points out that the sustainable development of the sports industry has encompassed aspects such as environmental governance, social inclusion, health promotion, and institutional accountability (Hautbois & Desbordes, 2023). When ESG is applied to sports center governance, E (Environmental) can correspond to energy saving, facility efficiency, and resource use. S (Socia) can correspond to health promotion, disadvantaged accessibility, and participation diversity; G (Governance) involves information disclosure, fair site allocation, operational accountability, and cross-departmental collaboration (Table 2). Recent ESG campaign governance

research also advocates that movement organizations need to establish actionable ESG implementation paths, rather than just making sustainability an image statement (Yu & Jeong, 2025).

Table 2: ESG is applied to ESG and observation indicators in sports centers

ESG	The meaning of sports center governance	Observable metrics
E	Energy saving, low-carbon venues, equipment maintenance, resource efficiency	energy use, carbon emissions, facility efficiency
S	Social inclusion, health promotion, access to the disadvantaged, and participation in diversity	participation rate, disadvantaged participation rate, satisfaction
G	Transparent governance, stakeholder collaboration, accountability, and resource allocation	Decision-making transparency, grievance response, and cooperation mechanisms

2.4 System dynamics and sports service governance

Systems dynamics is suitable for analyzing complex governance problems with multiple stakeholders, time delays, resource constraints, and feedback effects. Sterman et al. pointed out that system dynamics can help researchers understand cross-functional, nonlinear, and long-term cumulative management system behavior (Sterman et al., 2015). For example, increased participation may increase revenue and community trust, but if service capacity is not expanded simultaneously, it may also lead to congestion, decreased satisfaction, and user churn. Therefore, inclusive governance in sports centers should not be measured only by static performance indicators, but should be analyzed through causal loop diagrams and stock flow models to analyze the dynamic relationship between participation, capacity, trust, satisfaction, and resource reinvestment.

2.5 Literature gaps and analytical propositions

Based on the above literature, existing studies have discussed sports governance, social inclusion, sustainable sports services, and systems dynamics methods, but few have integrated the four into the dynamic model of inclusive governance in sports centers. In particular, existing research is biased towards institutional descriptions, governance principles, or performance evaluations, and rarely explains how inclusive governance is formed, strengthened, or declined through feedback loops over time. Based on this, this study proposes five analytical propositions: P1, participation in diversity enhancement will enhance community trust and public value; P2, insufficient service capacity will weaken the positive effect of participation growth on satisfaction; P3, improved governance transparency will promote user trust and continuous participation; P4, financial pressures may compress inclusive investment and improve exclusion risks; P5, adaptive governance can improve resource allocation through user feedback, thereby strengthening the sustainable service capabilities of sports centers. This integrated perspective can complement the lack of discussion of dynamic mechanisms and long-term policy effects in movement governance research (Madsen & Glebova, 2025; Shilbury et al., 2016;)。

3 Research methodology

3.1 Study design

This study adopts a conceptual system dynamics research design to construct a dynamic model of inclusive governance in the sports center, rather than a single cross-sectional causal test. System dynamics emphasizes the interaction between stock, flow, feedback loops, and time delays in social system behavior, and is suitable for analyzing the effects of multi-stakeholder, resource constraints, and policy interventions in public service governance (Forrester, 1961; Sterman, 2002). This study positions the sports center as a complex governance system that combines the government, outsourced operators, professional sports organizations, coaches and users. System dynamics is designed to understand complex system structures and dynamic behaviors, and has been widely used in the analysis of management and public policy issues.

3.2 Study scope and model boundaries

This study takes urban public sports centers as the analysis field, and the model boundaries cover three types of actors: first, the governance and operation side, including local governments, outsourcing operators, and venue managers; second, the service supply side, including professional sports organizations and individual coaches; Third, the user side, including teenagers, the elderly, beginners, regular exercisers and disadvantaged groups. The study focused on inclusive governance-related variables such as service capacity, participation

diversity, accessibility, satisfaction, trust, governance transparency, and resource reinvestment. The literature on sports governance points out that the governance of sports organizations has expanded from internal management to the integration of systems, stakeholders, and accountability, so this study regards the sports center as a multi-party collaborative governance system (Dowling et al., 2018;

Table 3: Model boundaries

category	Scope of inclusion
field	Urban public sports center
Service Providers	Outsourcing operators, professional sports organizations, and individual coaches
Users	Teenagers, the elderly, beginners, regular exercisers, and disadvantaged groups
Governance variables	Transparency, coordination, resource allocation, and feedback mechanisms
Operational variables	Service capacity, venue utilization, course supply, and financial reinvestment
Social variables	Engagement rate, participation diversity, satisfaction, trust, retention

3.3 System Dynamics Modeling Program

This study was conducted according to a five-stage modeling process: first, to define the problem scenario and illustrate the governance tension between operational efficiency, publicity and inclusive participation of sports centers; second, it proposes dynamic hypotheses to illustrate that a reinforcement or equilibrium loop may be formed between participation, trust, capacity, and resource reinvestment; third, establish a causal circuit diagram to identify the main positive and negative causal relationships; Fourth, establish a stock flow model and set inclusive participation, service capacity, trust capital and governance capabilities as core stocks; Fifth, conduct scenario simulations to compare the long-term effects of different governance strategies. Group modeling studies have also shown that system dynamics models can help stakeholders build consensus and promote organizational learning and policy commitment (Vennix, 1996).

3.4 Core variable design

The core stock of this study includes "inclusive participation stock", "service capacity stock", "trust capital stock", "governance capacity stock" and "financial resource stock". The main traffic includes new participant inflow, continuous participation flow, churn flow, capacity expansion flow, capacity depreciation flow, trust accumulation flow, and trust depletion flow. Secondary variables include service satisfaction, venue congestion, affordability, curriculum diversity, governance transparency, and inclusive investment. This design reflects that sports center governance is not a single performance variable, but an interplay between social participation, service quality, financial sustainability, and governance trust. The core variables are summarized in Table 4.

Table 4. Core Variables of Sports Center Inclusive Governance System Dynamics Model

Type	Variable name	Description
Stock	Inclusive Participation	The cumulative degree of participation of different ethnic groups
Stock	Service Capacity	Venues, courses, coaches and equipment can be serviced
Stock	Trust Capital	User trust in the governance and services of the sports center
Stock	Governance Capability	Organizational coordination, transparent decision-making, and responsiveness
Stock	Financial Resource Base	resources that can be invested in facilities, curriculum and inclusive services
Flow	Participation Inflow	New users join and engage with underprivileged groups
Flow	Retention Flow	User engagement continues
Flow	Dropout Flow	Loss due to crowding, dissatisfaction, or cost pressure
Auxiliary	Satisfaction	feelings about service quality, accessibility, and fairness
Auxiliary	Congestion	Service pressure created by demand exceeding capacity
Auxiliary	Inclusion Investment	barrier-free facilities, preferential courses, diversified curriculum design, etc

3.5 Scenario simulation design

This study plans five sets of scenarios: S0 (Baseline) is the baseline scenario, assuming that the existing capacity and governance model are maintained; S1 (Inclusive Program Design) designs scenarios for inclusive courses, adding courses that are friendly to the elderly, beginners, and disadvantaged groups; S2 (Capacity

Allocation Policy) is a capacity allocation scenario that adjusts venues and course resources according to peak and off-peak needs and group characteristics. S3 (Transparent Adaptive Governance) aims to transparently adapt to governance scenarios, strengthen information disclosure, user feedback, and stakeholder coordination; S4 (Cost-Pressure Scenario) is a cost-pressure scenario that simulates a decline in inclusive investment due to increased financial pressure. Scenario simulations will compare the long-term impact of each strategy on engagement diversity, service satisfaction, trust capital, and sustainable operations.

Table 5: Scenario simulation design

Scenario	Description	Expected effect
S0	Maintain existing governance and capacity allocation	Participation grows slowly and may be limited by capacity
S1	Increase senior, beginner, disadvantaged and family-friendly courses	Enhance participation diversity and social value
S2	Adjust capacity according to peak off-peak, group needs and site utilization	Reduce crowding and increase satisfaction
S3	Strengthen information disclosure, user feedback, and stakeholder coordination	Increase trust, retention, and re-engagement
S4	Financial pressure is rising and inclusive investment is declining	It may lead to the exclusion of vulnerabilities and a decline in trust

Model checks are based on structural rationality, variable unit consistency, extreme condition testing, and behavior pattern checking. The simulation results are not intended to predict a single value, but to compare the trends of system behavior arising from different governance strategies. Lane (2000) points out that system dynamics is not a deterministic prophetic tool, but rather a policy learning method that helps understand how structures partially shape behavior. Therefore, this study uses scenario simulation as a governance strategy comparison and management learning tool to identify possible positive loops and unexpected exclusion effects.

4 Research results

4.1 Research Framework

The results of this study present a "dynamic model of the inclusive governance system of sports centers" (Figure 1). This model takes ESG as the upper governance principle, linking public sports services, participation diversity, service capacity, user trust, and resource reinvestment. The sports center is not just a venue, but a service ecosystem composed of the government, outsourced operators, professional sports organizations, coaches and users, involving multiple operations and inclusive governance. Sports participation research also pointed out that there is a cross-sectoral linkage between public policy, facility supply, and user health, well-being, and social capital, so it is suitable for a systematic perspective (Kumar etal., 2018).

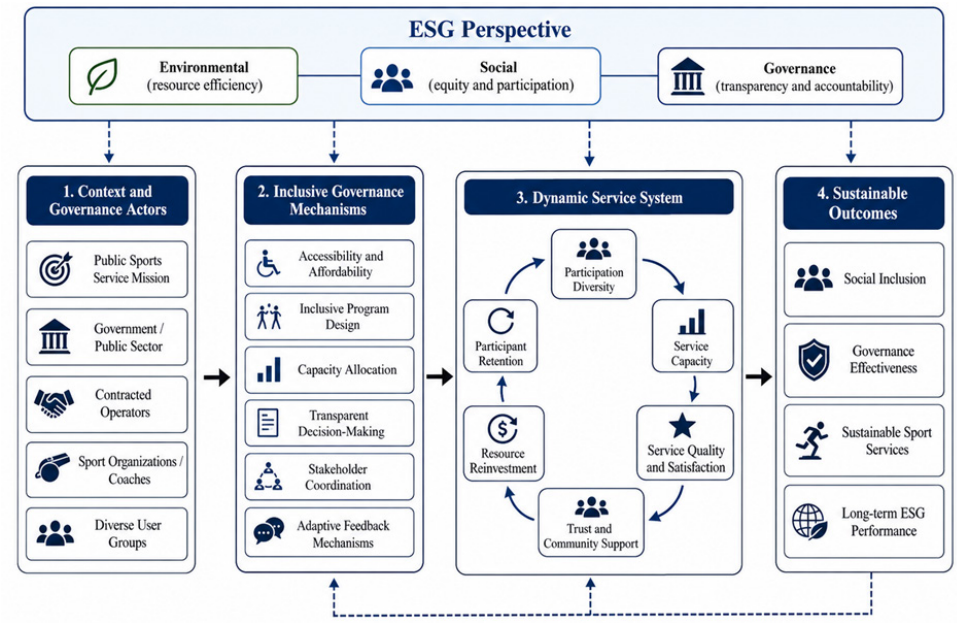


Figure 1: Research Framework of Inclusive Governance of Sports Centers

4.2 Causal circuit analysis

The causal loop diagram shows that the model in this study (Figure 2) contains three reinforcement loops and two equilibrium loops, which are summarized in Table 6。 R1 "Participation-Trust Enhancement Loop" explains that increasing participation diversity can increase community recognition and public trust, thereby promoting re-participation. R2 "Service Quality - Retention Enhancement Loop" shows that improved service quality will enhance satisfaction, retention rate, and stable revenue, allowing sports centers to have more resources to reinvest in courses and facilities. R3 "Governance Transparency - Stakeholder Support Strengthening Loop" points out that information disclosure and feedback mechanisms can enhance trust and collaboration capabilities. In contrast, the B1 "congestion-service dilution balance loop" shows that if the increase in participation exceeds capacity, it will lead to congestion, decreased satisfaction, and churn. B2 "Cost Pressure-Inclusive Investment Restriction Loop" pointed out that financial pressure may compress public welfare courses and support for the disadvantaged. This result echoes the emphasis placed in the movement governance literature on institutions, accountability, and stakeholder interaction (Dowling et al., 2018; Shilbury et al., 2016) .

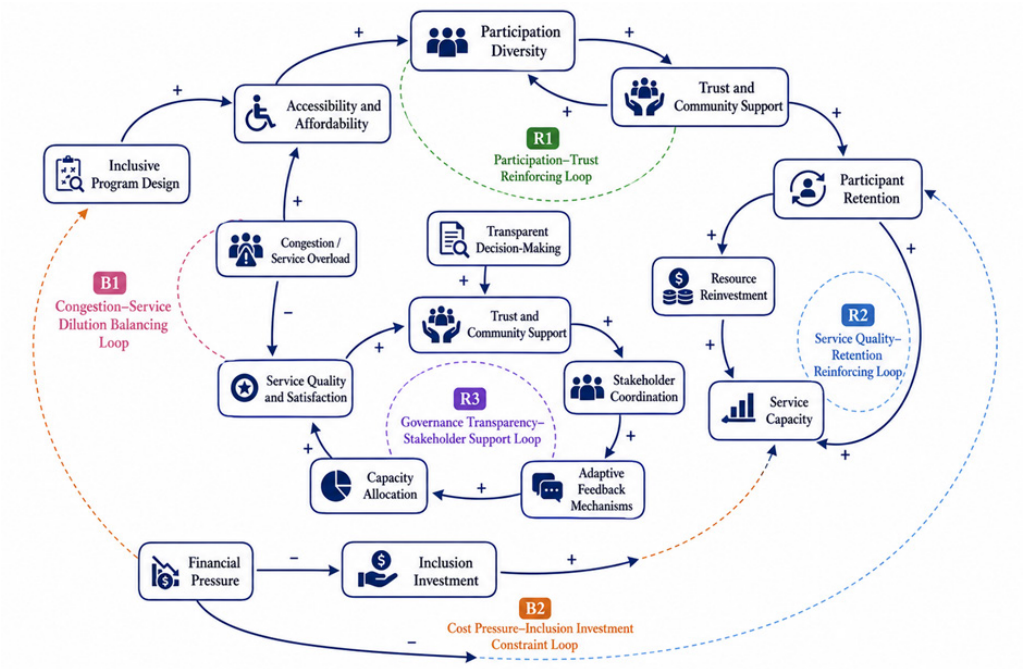


Figure 2: Causal Loop Diagram of Sports Center Inclusive Governance

Table 6: Type of Causal Loop Diagram on Sports Center Inclusive Governance

circuit	Type	Description
R1 Participation–Trust Reinforcing Loop	Strengthen the circuit	Increased participation → Increased community identity → Increased trust → increased re-participation
R2 Service Quality–Retention Reinforcing Loop	Strengthen the circuit	Service Quality Improvement → Satisfaction Improvement → Retention Rate Improvement → Stable Revenue and Reinvestment
R3 Governance Transparency–Stakeholder Support Loop	Strengthen the circuit	Increased transparency → Increased trust → Stakeholder support → Improved governance capabilities
B1 Congestion–Service Dilution Balancing Loop	Balance circuit	Increased participation → increased congestion → decreased satisfaction → inhibited participation growth
B2 Cost Pressure–Inclusion Investment Constraint Loop	Balance circuit	Increased financial pressure → Decline in inclusive investment → Decrease in disadvantaged participation → Decrease in public value

4.3 Stock flow model results

The stock flow model (Figure 3) sets "inclusive participation stock", "service capacity stock", "trust capital stock", "governance capacity stock" and "financial resource stock" as core stock. The model results show that inclusive participation is not only affected by the number of courses, but also by service capacity, affordability, course adaptation, governance transparency, and trust accumulation. When service capacity and governance capabilities increase simultaneously, participation inflow and retention flow will increase; Conversely, capacity depreciation, crowded venues, or opaque decision-making can increase churn and trust erosion. System dynamics emphasizes that the behavior of complex service systems often stems from feedback, latency, and cross-functional interactions rather than single-variable effects (Sterman et al., 2015; Sterman, 2002)。

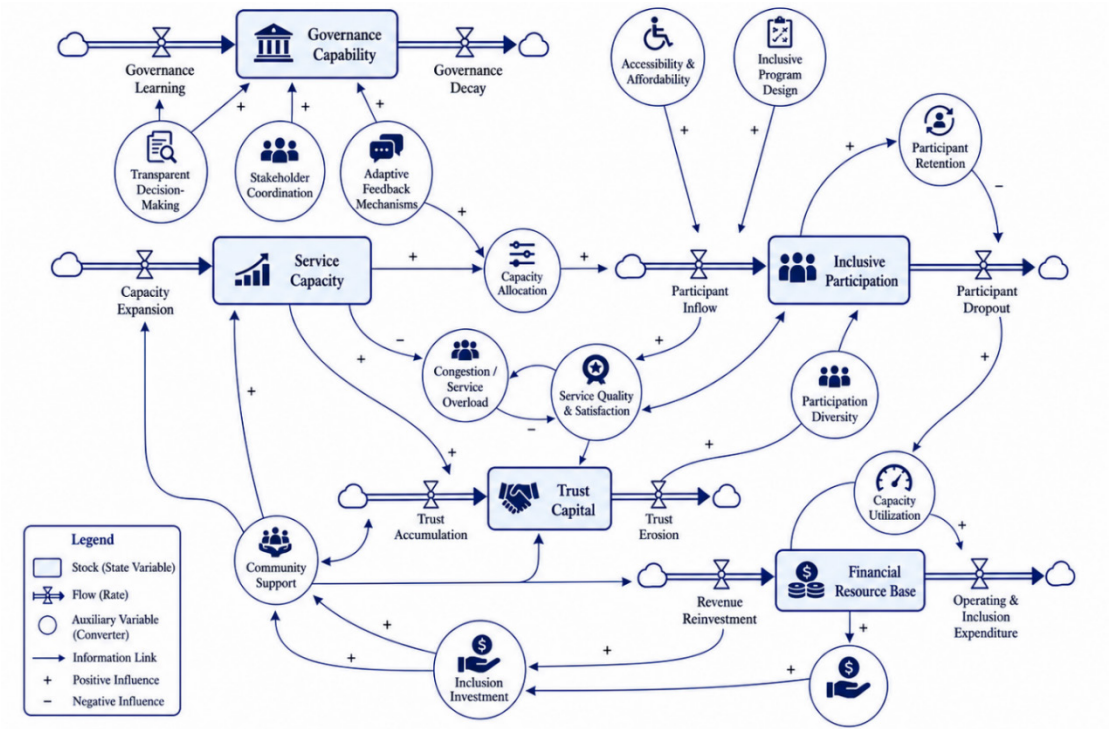


Figure 3: Stock-and-Flow Model of Inclusive Sports Center Governance

4.4 Scenario simulation results

Scenario simulations compare five governance strategies. In the S0(baseline scenario), the number of participants increases initially, but slows down later due to capacity constraints. The S1(Inclusive Program Design) is designed to increase the participation of the elderly, beginners, and disadvantaged groups, but if the capacity is not adjusted simultaneously, satisfaction may decrease. S2(Capacity Allocation) capacity configuration scenarios reduce peak congestion and improve service quality and retention. S3(Transparent Adaptive Governance) Transparently adapts to governance scenarios is the most stable for trust capital and long-term engagement, as it can continuously modify resource allocations through user feedback. The S4(Cost Pressure) cost stress scenario shows negative outcomes: inclusive investment declines, disadvantaged participation decreases, and public trust decreases. This result (Table 6) also echoes the view of social inclusion research: sports participation does not necessarily lead to inclusion, but must be supported through theoretical orientation and institutionalized design (Coalter, 2017).

Table 6: Scenario Simulation Results of Inclusive Governance Strategies

Scenario	Overall Interpretation
S0	Under the existing governance model, various indicators grew slowly, but in the later stage, due to insufficient service capacity and governance adjustment, the overall performance tended to flatten.
S1	Inclusive curriculum design is the best way to increase participation diversity, but if capacity allocation is not improved simultaneously, service satisfaction may be limited by congestion.
S2	Capacity allocation strategies can effectively reduce congestion, improve service satisfaction and retention, and significantly contribute to operational stability.

Scenario	Overall Interpretation
S3	Transparent adaptive governance performs best overall, strengthening engagement, trust, service quality, and long-term ESG governance performance at the same time.
S4	Cost stress scenarios have led to a decline in inclusive investment, worsening participation diversity, satisfaction, and trust capital.

4.5 Summary of key findings

This study summarizes five key findings: First, the key to inclusive governance in sports centers is not simply to increase utilization, but to balance participation growth with service capacity. Second, trust capital is the core intermediary mechanism that connects engagement, satisfaction, and long-term retention. Third, governance transparency and user feedback can form adaptive governance capabilities. Fourth, if financial pressure takes precedence over public value, it will increase the risk of eliminating the vulnerable. Fifth, systems dynamics can integrate ESG, inclusive governance, and sports service management into a simulated decision support framework.

5 Discussion

5.1 Theoretical implications

This study regards sports centers as a "public sports service governance system" rather than a single venue management unit, supplementing the existing sports governance research that focuses on system, board of directors, or organizational performance. Movement governance research has pointed out that governance issues involve institutional arrangements, accountability, stakeholder relations, and power allocation (Dowling et al., 2018). Collaborative governance research also emphasizes the need for movement organizations to address complex public issues through cross-organizational coordination (Shilbury et al., 2016). Therefore, the theoretical implications of this study are to show that inclusive governance is not a static norm, but a dynamic process formed by participation, trust, capacity, and resource reinvestment.

5.2 ESG and system dynamics integration

This study translates ESG into a modelable governance structure: E corresponds to facility energy efficiency and resource allocation, S corresponds to participation diversity, health promotion, and social inclusion, and G corresponds to transparent decision-making, accountability, and stakeholder alignment. System dynamics emphasizes that the behavior of complex service systems often comes from feedback, delays, and cross-departmental interactions, rather than a single factor linear effect (Sterman et al., 2015). Therefore, this study uses a causal loop and stock flow model to transform ESG from an evaluation indicator into a dynamic governance mechanism.

5.3 Management implications

For sports center managers, the results show that increasing utilization does not necessarily mean governance success. If service capacity, venue configuration, and course design are not synchronized, increased participation can lead to congestion, decreased satisfaction, and user churn. Social inclusion research also points out that sports participation requires theoretical orientation and institutional support, otherwise it is difficult to naturally produce fairness and inclusion effects (Coalter, 2017). Therefore, managers should establish user feedback mechanisms, off-peak capacity adjustments, disadvantage-friendly courses, and transparent information disclosure systems.

5.4 Policy implications

For policymakers, outsourcing performance of sports centers should not be measured solely by revenue, attendance, or venue utilization, but should include indicators such as participation diversity, disadvantaged accessibility, service satisfaction, governance transparency, and community trust. This study pointed out that public sports centers should be positioned as health promotion and social inclusion infrastructure, rather than simply operating venues. This perspective also echoes the positioning of sports centers as platforms for public health, social inclusion, and community well-being.

6 Conclusion

6.1 Research conclusions

This study constructs an inclusive governance model for sports centers from the perspective of system dynamics, showing that sports centers are not only service units that provide courses and venues, but also a public sports service system composed of the government, outsourcing operators, professional organizations, coaches, and multiple users. This positioning positions the Sports Center as a platform for public health, social inclusion, and community well-being. The results point out that the key to inclusive governance lies in the dynamic balance between participation diversity, service capacity, trust capital, governance transparency, and resource reinvestment. If the governance mechanism can strengthen user feedback, capacity allocation, and disadvantaged accessibility, it will form a positive cycle of "participation-trust-re-investment"; Conversely, excessive pursuit of usage and short-term revenue can lead to congestion, decreased satisfaction, and exclusion. This finding echoes the emphasis placed on stakeholder coordination and accountability in movement governance research, and is also in line with the analytical perspective of system dynamics on feedback, delay, and cross-departmental interactions (Dowling et al., 2018; Shilbury et al., 2016; Sterman et al., 2015).

6.2 Study limitations

This study is a conceptual system dynamic model and has not yet used actual sports center operation data for parameter calibration. Different cities, outsourcing systems, charging models, and user structures may also affect the applicability of the model. In addition, this article mainly discusses the overall mechanism of inclusive governance, and there is still room for deepening the barriers to participation of subgroups such as the elderly, people with disabilities, low-income groups, or female users. Research on social inclusion in sports also reminds that sports participation does not necessarily produce natural inclusive effects and still needs to be verified through theory-oriented evaluation and institutionalized design (Coalter, 2017).

6.3 Future research suggests

Future research can collect data on the flow of people in sports centers, course registrations, satisfaction, complaints, charges, and energy usage to calibrate the stock flow model and conduct scenario simulations. Secondly, it can be combined with interviews, questionnaires, and group modeling to incorporate the perspectives of governments, operators, coaches, and users. Finally, the ESG Inclusive Governance Dashboard can be further developed as a decision-making support tool for public value, operational sustainability, and social inclusion in sports centers.

7 Ethics Statement

This study does not address research ethics issues.

8 Artificial intelligence declares

AI tools are only used for language retouching, structure assistance, and editing support. All content, arguments, citations, and conclusions are reviewed, verified, and fully responsible for the author.

References

Coalter, F. (2017). Sport and social inclusion: evidence-based policy and practice. Social inclusion, 5(2), 141-149.

Dowling, M., Leopkey, B., & Smith, L. (2018). Governance in sport: A scoping review. Journal of sport management, 32(5), 438-451.

Forrester, J.W. (1961) Industrial Dynamics. MIT Press, Cambridge, Mass.

Hautbois, C., & Desbordes, M. (2023). Sustainability in Sport: Sport, Part of the Problem… and of the Solution. Sustainability, 15(15), 11820.

Jeanes, R., Spaaij, R., Farquharson, K., McGrath, G., Magee, J., Lusher, D., & Gorman, S. (2021). Gender relations, gender equity, and community sports spaces. Journal of Sport and Social Issues, 45(6), 545-567.

Kumar, H., Manoli, A. E., Hodgkinson, I. R., & Downward, P. (2018). Sport participation: From policy, through facilities, to users' health, well-being, and social capital. Sport management review, 21(5), 549-562.

Lane, D. C. (2000). Should system dynamics be described as a 'hard'or 'deterministic'systems approach?. Systems research and behavioral science: the official journal of the international federation for systems research, 17(1), 3-22.

Madsen, D. Ø., & Glebova, E. (2025). Sports Industry 5.0: reimagining sport through technology, humanity, and sustainability. Frontiers in Sports and Active Living, 7, 1640362.

Shilbury, D., O'Boyle, I., & Ferkins, L. (2016). Towards a research agenda in collaborative sport governance. Sport management review, 19(5), 479-491.
Sterman, J. D. (2002). Business Dynamics, Systems Thinking and Modeling for a Complex World. Massachusetts Institute of Technology Sloan School of Management.
Sterman, J., Oliva, R., Linderman, K., & Bendoly, E. (2015). System dynamics perspectives and modeling opportunities for research in operations management. Journal of Operations Management, 39, 1-5.
Vennix, J. A. M. (1996). Group model building: Facilitating team learning using system dynamics. John Wiley & Sons.
Yu, J. G., & Jeong, Y. (2025). Enhancing Environmental, Social, and Governance Practices in Professional Sports Clubs: A Delphi Study Approach. Sage Open, 15(4), 21582440251395587.

Towards Smart and Sustainable Preschools: System Dynamics Modeling of AI-Empowered Educational Governance under ESG Principles

Chia-Fen Lin[1], Yin-Feng Chen[2], Rou-Ya Lin[1]
[1] Department of Early Childhood and Family Education, National Taipei University of Education, Taipei, Taiwan
[2] Department of Business Administration, Shih Chien University. Taipei, Taiwan

Abstract

This study aims to construct a smart and sustainable preschool governance model based on Environmental, Social, and Governance (ESG) principles by integrating artificial intelligence (AI)-enabled decision support and system dynamics (SD). Faced with multiple challenges such as declining birth rates, teacher stability, educational quality assurance, resource constraints, and sustainability expectations, preschools are no longer merely care and educational settings but also important early learning organizations that promote educational equity, organizational resilience, and sustainable development. This study adopts conceptual modeling and scenario-based analysis. First, ESG principles are translated into core governance variables, including environmental efficiency, social equity, teacher support, parental trust, and governance transparency. Second, causal loop diagrams (CLDs) are used to construct reinforcing and balancing loops related to AI data learning, teacher quality, governance transparency, cost pressure, and teacher workload. Third, a stock-and-flow model (SFD) is developed to explain the long-term accumulation mechanisms among educational quality, AI governance capability, ESG performance, and institutional resilience. The scenario analysis suggests that the integrated AI-enabled ESG governance scenario is expected to generate the strongest long-term improvements. The research results show that AI should not be seen as a tool to replace the education profession, but should be used as an important foundation to support knowledge management, policy simulation, and sustainable governance decision-making in preschools.

Keywords: ESG governance, artificial intelligence, system dynamics, smart preschool, sustainable education governance

Paper type: Academic Research Paper

1 Introduction

As an important field for early education and family support, preschools are facing multiple pressures such as improving education quality, the impact of declining birth rates, teacher stability, resource constraints, and sustainable governance requirements. Heckman (2006) pointed out that early ability formation has a dynamic cumulative effect, and early investment will affect subsequent educational and social outcomes. Nores and Barnett (2010) also found that early education interventions can deliver long-term benefits of cognitive, behavioral, health, and learning outcomes in different country contexts. This perspective suggests that preschool governance should not be viewed as a mere administrative issue, but should be understood as a public governance issue related to educational equity, social resilience, and sustainable development.

In recent years, sustainability and Environmental, Social, and Governance (ESG) principles have gradually expanded from corporate governance to educational organizations. Early sustainability education research indicates that preschools can cultivate children's initial understanding of nature, social responsibility, and sustainable living through games, life practices, and environmental exploration (Engdahl & Furu, 2022). Borg and Samuelsson (2022) also emphasize that young children are not merely passive recipients of sustainability education; rather, they can demonstrate agency as active learning subjects through participation and action. Recent eco-program research in preschools further points out that sustainable early childhood education should integrate ecological, social, and economic aspects rather than being limited to environmental education activities (Vodopivec & Šindic, 2025).

Existing research mostly focuses on sustainable education curricula, early childhood learning activities, or teacher attitudes, and there is still a lack of systematic models for how ESG principles translate into preschool governance indicators, resource allocation mechanisms, and long-term performance management. Although the Korean daycare ESG study found a positive correlation between ESG management and teachers' awareness of social issues and environmentally friendly education attitudes, it still favored cross-sectional empirical

evidence and less dealt with the time delay and feedback relationship between governance variables (Lim & Kim, 2025).

The application of artificial intelligence (AI) in education is expanding rapidly, and relevant research points out that AI can support personalized learning, predictive analytics, and educational decision-making, but it also involves data ethics, algorithmic bias, and governance accountability issues (Strielkowski et al., 2025; Zawacki-Richter et al., 2019). Therefore, AI should not be simplified as a teaching tool in preschools, but should be positioned as a knowledge management foundation to assist preschools in data analysis, policy simulation, and governance decision-making.

Based on the above gaps, this study extends the previous research positioning by integrating ESG governance, AI-enabled decision support, and system dynamics modeling into a smart and sustainable preschool governance framework. System dynamics is suitable for analyzing policy issues with time delays, nonlinearity, and feedback loops in education systems, and can help uncover the dynamic relationship between teacher quality, child development, resource allocation, parental trust, and governance transparency (Groff, 2013). Therefore, the purpose of this study is: first, to construct the core variables of ESG education governance in preschools; second, causal loop diagrams (CLDs) and a stock-and-flow model (SFD) are used to present the governance feedback structure; Third, AI-supported ESG policy scenarios are analyzed to assess their impact on education quality, organizational resilience, and sustainability performance.

The main questions of this study are: 1. What are the core ESG variables of the smart sustainable preschool governance system? 2. How does AI-supported data analysis improve the quality of preschool governance decision-making? 3. What reinforcing and balancing loops exist among teacher quality, parental trust, resource allocation, and child development? 4. What are the differences between different ESG policy scenarios on the long-term governance performance of preschools?

2 Literature review

2.1 ESG principles and educational governance

ESG principles were originally used to evaluate an organization's sustainability performance in environmental, social, and governance aspects, but in recent years, they have gradually extended to the field of education governance. For educational institutions, Environmental can be transformed into campus resource efficiency, energy conservation and carbon reduction, and environmental education. Social can connect educational equity, child well-being, teacher support and family participation; Governance involves transparent decision-making, budget accountability, data governance, and stakeholder communication. Higher education research points out that ESG implementation should not be limited to curriculum or image claims, but should be included in organizational governance, operation management, and sustainability performance evaluation. Research by childcare institutions also shows that ESG management can enhance the awareness of education and care workers about social issues and environmentally friendly education. Therefore, preschools can be regarded as sustainable education organizations with public responsibility, rather than mere care service units (Rosli & Azmi, 2025; Finatto et al., 2024).

Table 1: preschool governance under the ESG dimension

ESG aspect	The meaning of preschool governance	Possible variables
Environmental	Energy conservation, carbon reduction, green campus, environmental education	energy use, water resource use, and investment in green facilities
Social	Educational Equity, Child Well-being, Teacher Support, Family Engagement	affordability, disadvantaged enrollment rate, teacher satisfaction, child development
Governance	Transparent governance, data ethics, financial accountability, parental communication	Budget transparency, data governance maturity, parental trust

2.2 Sustainable Early Childhood Education and Early Learning Organization

Sustainable early childhood education emphasizes the formation of sustainable values in children's daily lives, games, environmental exploration, and social interactions. Engdahl and Furu (2022) pointed out that early childhood education is suitable as an early stage for sustainable social transformation due to its holistic, life-oriented, and child-centered nature. Borg and Samuelsson (2022) further emphasize that young children are not passive recipients of sustainable education but can demonstrate agency through participation, expression, and action. It can be seen that if preschools want to become early learning organizations, they must pay attention

to child development, teacher professionalism, family participation, and organizational learning capabilities, and upgrade sustainability from curriculum activities to organizational governance logic.

2.3 AI-powered educational governance and knowledge management

The application of AI in education has gradually expanded from personalized learning, learning analysis, and risk prediction to administrative management and governance decision-making. A systematic review by Zawacki-Richter et al. (2019) pointed out that AI education research has covered intelligent teaching systems, assessment, prediction, and support services, but it still needs to pay attention to the role of educators and ethical issues. Strielkowski et al. (2025) also point out that AI-driven adaptive learning can facilitate educational transformation, but it must be integrated with sustainability goals, accessibility, and equity. From the perspective of knowledge management, Alavi and Leidner (2001) regard information technology as an important foundation to support the creation, storage, transfer, and application of knowledge in organizations. Recent AI and knowledge management research reminds that while AI can improve knowledge processing capabilities, it also poses challenges in data quality, trust, ethics, and organizational implementation (Rezaei, 2025). Therefore, this study positions AI as a decision support and knowledge transformation tool for preschool governance, rather than an automated mechanism that replaces professional judgment by teachers and principals.

2.4 Application of System Dynamics in Educational Policy and Governance Research

Educational governance has the characteristics of multiple feedbacks, time delays, and policy trade-offs, making it suitable for system dynamics analysis. Groff (2013) pointed out that system dynamics models can help education policy researchers identify hidden system structures and simulate future policy scenarios. Kennedy (2011)'s review of the dynamic model of education policy systems also shows that system dynamics (SD) has been used for planning, resource allocation, and policy evaluation at the school, regional, and tertiary levels. Pedamallu et al. (2010) used primary education enrollment as an example to illustrate the interaction between infrastructure, teaching quality, sense of community, and educational accessibility. Therefore, the long-term dynamics of teacher quality, enrollment stability, parental trust, financial resources, and sustainable investment in preschool governance can also be presented through causal loop diagrams and stock-and-flow models.

2.5 Literature gaps and research framework derivation

In the literature review , the introduction of ESG education, sustainable early childhood education, AI education applications, and dynamic models of education systems have been discussed, but there are still three gaps: first, the indicatorization and systematic application of ESG in preschool governance are still insufficient; Second, AI education research mostly focuses on learning or teaching, and rarely explores its role in preschool governance, resource allocation, and accountability. Third, existing research on sustainability in early childhood education rarely deals with dynamic feedback between teacher quality, child development, parental trust, governance transparency, and resource constraints. Based on this, this study extends the aforementioned research positioning and constructs an integrated framework of "ESG Principles-AI Decision Support-System Dynamic Simulation" to analyze the governance mechanism and policy scenarios of smart and sustainable preschools.

3 Research methodology

3.1 Study design

This study adopts conceptual modeling and scenario simulation research design, using system dynamics as the main method to construct an AI-supported ESG preschool education governance model. System dynamics is suitable for dealing with common feedback loops, time delays, and policy trade-off problems in education governance. Groff (2013) pointed out that system dynamics models can be used in education policy design and simulation to help researchers understand the long-term behavior of complex education systems. Sterman (2000) also emphasizes that SD modeling can analyze complex organizational systems through problem definition, dynamic assumptions, model building, testing, and policy design.

ESG defines the normative governance goals, AI provides data-driven analytical and decision-support capabilities, and system dynamics offers a modeling approach for understanding feedback structures, time delays, and policy trade-offs in preschool governance.

3.2 System boundaries and units of analysis

In this study, the term "preschool" refers broadly to early childhood education and care institutions serving young children before primary schooling, including kindergartens and daycare centers where applicable. The unit of analysis is the preschool institution. Preschool principals are considered key governance actors, while the system boundaries cover educational quality, teacher professional capacity, child development, parental trust, resource allocation, environmental efficiency, governance transparency, and AI decision-support capability.

3.3 Core variable design

The core variables are organized into several categories: first, education quality, including teachers' professional ability, teacher-student interaction quality, and child development outcomes; Early education research shows that teacher professional development and teaching quality affect early childhood learning and social-emotional development (Brunsek et al., 2020). Second, social equity, including affordability, access to preschool for disadvantaged children, and family support. Third, parental trust, including home communication, information transparency and parental participation; Family engagement research points out that parent-teacher relationships have a significant impact on early education outcomes (Otero-Mayer et al., 2025). Fourth, environmental performance, including energy, water resources, and green campus investment. Fifth, governance transparency, including budget disclosure, performance disclosure, and data ethics. Sixth, AI governance capabilities, including predictive analysis, resource allocation suggestions, risk warnings, and scenario simulations. These variables are organized in Table 2.

Table 2: Core Variables of AI-empowered ESG Preschool Governance Model

category	variable	Operational definitions	ESG correspondence
Quality of education	Teacher Quality	Teachers' professional ability, stability and teaching quality	S
Child development	Child Development Outcomes	Cognitive, social-emotional and learning readiness	S
Parental trust	Parental Trust	Parents' trust in the quality, transparency, and safety of the preschool	S/G
Governance capabilities	Governance Transparency	The degree of disclosure of budget, quality, and decision-making information	G
AI capabilities	AI Decision Support Capability	Data analysis, forecasting, simulation, and decision support capabilities	G
Environmental performance	Environmental Efficiency	energy, water resources, and campus environmental efficiency	E
Financial resources	Funding Availability	income, grants, investments, and working capital	G
Accessibility	Access Equity	Disadvantaged families, affordability, and access to preschool	S

3.4 Causal Loop Diagram (CLD) Design

This study establishes a causal loop diagram (CLD) to present the reinforcing and balancing loops in the preschool governance system. Three reinforcing loops are identified: R1, the "AI Data Learning–Governance Quality Reinforcing Loop," illustrates how AI analytics capability improves decision quality and resource allocation efficiency; R2, the "Teacher Quality–Child Development Reinforcing Loop," explains how teacher professional support promotes child development and parental trust; and R3, the "Governance Transparency–Parental Trust Reinforcing Loop," shows how transparent governance strengthens parental trust and feedback data quality. In addition, two balancing loops are identified: B1, the "Cost Pressure–Sustainability Investment Constraint Balancing Loop," and B2, the "Teacher Workload–Educational Quality Regulation Balancing Loop," which reveal the policy trade-offs between ESG investment and teacher workload pressure.

3.5 Stock-and-Flow Model Design

The stock-and-flow model (SFD) defines core stocks as Teacher Professional Capacity, Educational Quality Capital, Parental Trust Capital, AI Governance Capability, ESG Performance Stock, and Institutional Resilience. Corresponding flows include teacher professional growth rate, education quality improvement rate, trust accumulation rate, AI capability building rate, ESG investment rate, and resilience improvement rate. Groff

(2013) points out that stocks represent accumulated resources or capabilities in the system, while flows represent their rate of change; both can be used to simulate the behavior of the education system.

3.6 Scenario simulation design

This study plans five scenarios (Table 3): S0 baseline scenario; S1 Green Campus and Energy Saving Investment; S2 Educational Equity & Inclusion Enrollment; S3 Budget transparency and communication with parents; Integrated ESG governance powered by S4 AI. Pedamallu et al. (2010) used SD models to analyze the relationship between basic education enrollment and infrastructure, showing that scenario simulation is useful for evaluating education policy effectiveness. In this study, AI serves as a decision support tool for predicting needs, identifying risks, and comparing policy scenarios. Relevant studies also point out that AI education applications need to take into account human-centered control, stakeholder participation, and trusted governance (Zawacki-Richter et al., 2019; Alfredo et al., 2024).

To enhance model credibility, this study applies structural validation by comparing the proposed feedback loops with prior literature on early childhood education quality, ESG-oriented educational governance, AI-enabled education, and system dynamics modeling. Since this is a conceptual modeling study, parameter calibration and empirical validation are reserved for future research.

Since this study is positioned as a conceptual system dynamics modeling study, the scenario-based analysis is qualitative and theory-driven rather than parameter-based numerical simulation.

Table 3. Scenario Design for ESG-driven Preschool Governance Simulation

Situation	Policy setting	Main interventions	Expected effect
S0 Baseline	Maintain the current governance model	No additional AI or ESG investments	The quality of education is growing slowly, but it is susceptible to cost and teacher workload
S1 Eco-efficient Investment	Strengthen energy-saving and green facilities	environmental equipment, energy-saving management	Reduce operating costs and improve environmental performance over time
S2 Inclusive Enrollment Strategy	Strengthen educational equity	Subsidies for the disadvantaged, affordable programs	Improve social performance, but increase financial pressure in the short term
S3 Transparent Budget Governance	Strengthen governance transparency	Budget disclosure, parent communication platform	Enhance parental trust and engagement
S4 AI-enabled Integrated Governance	Integrate AI, ESG, and SD simulation	Predictive analytics, real-time dashboards, policy simulations	Improve overall resilience, quality, and governance accountability

4 Research results

4.1 ESG preschool governance Framework Results

The results of this study first establish a three-tier governance framework of "ESG Principles-AI Decision Support-System Dynamic Simulation" (Figure 1). Environmental is oriented towards energy, water resources, green campuses, and environmental education; The Social aspect corresponds to educational equity, child well-being, teacher support and family participation; The governance aspect corresponds to budget transparency, data ethics, performance disclosure, and parental trust. This framework continues the three core pillars proposed in the previous summary: ESG-aligned governance indicators, system dynamics feedback structures, and AI-driven data analytics (Finatto et al., 2024). Preschools are thus defined as an early learning organization, and their governance performance depends not only on administrative efficiency but also on the long-term accumulation of educational quality, social equity, and sustainability. Engdahl and Furu (2022) pointed out that early childhood education is an important field for sustainable social transformation. ESG education research also shows that the implementation of ESG in educational institutions should cover curriculum, operations, governance, and performance evaluation.

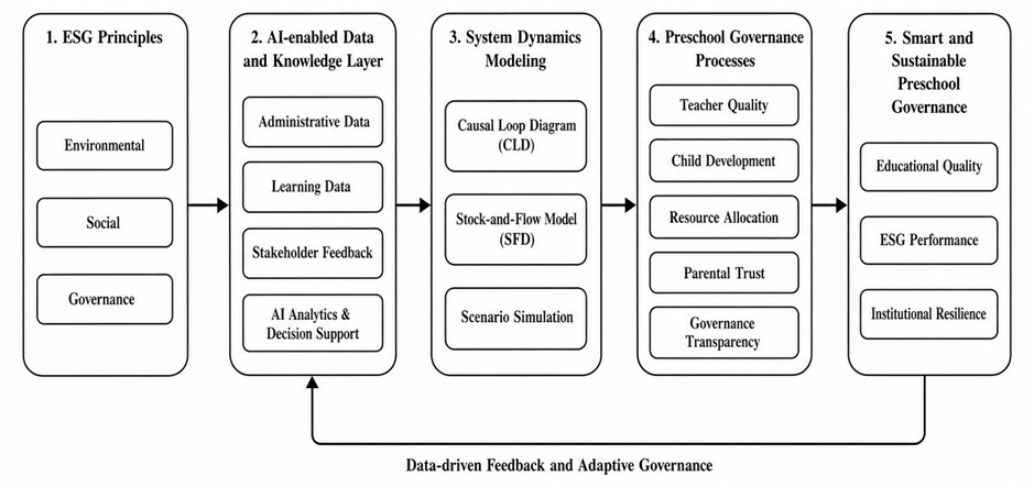


Figure 1: Research Framework of AI-empowered ESG Preschool Governance

4.2 Causal loop analysis results

The results of causal loop analysis (Figure 2) show that smart sustainable preschool governance includes at least three reinforcing loops and two balancing loops. R1, the " AI Data Learning–Governance Quality Reinforcing Loop " loop points out that after improving AI data analysis capabilities, it can improve enrollment prediction, teacher allocation, resource allocation, and risk warning, thereby improving governance quality and parental trust. The R2 " Teacher Quality–Child Development Reinforcing Loop " shows that teachers' professional support improves teaching quality, further promoting child development outcomes and parent satisfaction. R3 the " Governance Transparency–Parental Trust Reinforcing Loop " explains that budget and performance transparency can strengthen parental trust, increase parental feedback and participation, and thereby improve governance adaptability. This result echoes Alavi and Leidner's (2001) view of knowledge management systems: information technology can support knowledge creation, storage, transfer, and application.

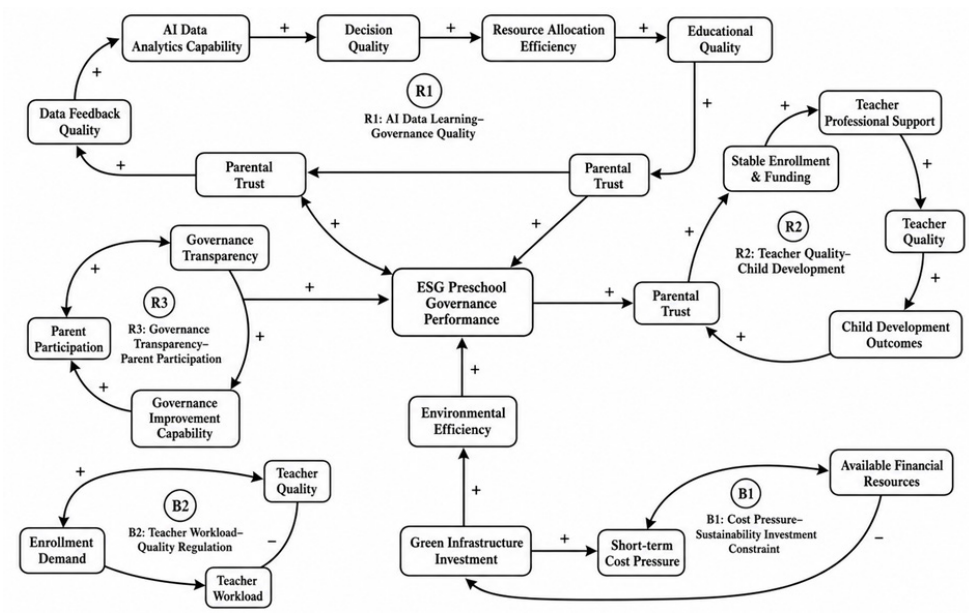


Figure 2: Causal Loop Diagram of AI-empowered ESG Preschool Governance System

At the same time, the model also reveals two restrictive structures. B1 " Cost Pressure–Sustainability Investment Constraint Balancing Loop" explains that although investment in green facilities and AI systems can improve long-term benefits, it may compress financial resources in the short term. B2 " Teacher Workload–Educational Quality Regulation Balancing Loop " pointed out that if the increase in enrollment is not synchronized with the expansion of teachers and support systems, it will increase teacher pressure and reduce the quality of education. A meta-analysis by Brunsek et al. (2020) pointed out that there is an important

association between the professional development of early childhood educators and children's outcomes, indicating that teacher competence is a key lever variable in the model.

4.3 Stock-and-Flow Model Results

The stock-and-flow model (Figure 3) decomposes preschool governance capabilities into six main stocks: Teacher Professional Capacity, Educational Quality Capital, Parental Trust Capital, AI Governance Capability, and ESG Performance Stock and Institutional Resilience. The corresponding flows include teacher professional growth rate, education quality improvement rate, trust accumulation rate, AI capability construction rate, ESG investment rate, and resilience improvement rate. This model shows that educational quality and parental trust cannot be formed quickly through short-term investment, but are gradually accumulated through teacher training, transparent governance, data feedback, and stable interaction. Groff (2013) pointed out that the system dynamics model helps to reveal hidden structures and long-term scenarios in the education policy system. Pedamallu et al. (2010) also pointed out that the quality and accessibility of basic education are influenced by the interaction of infrastructure, community conditions, and policy factors.

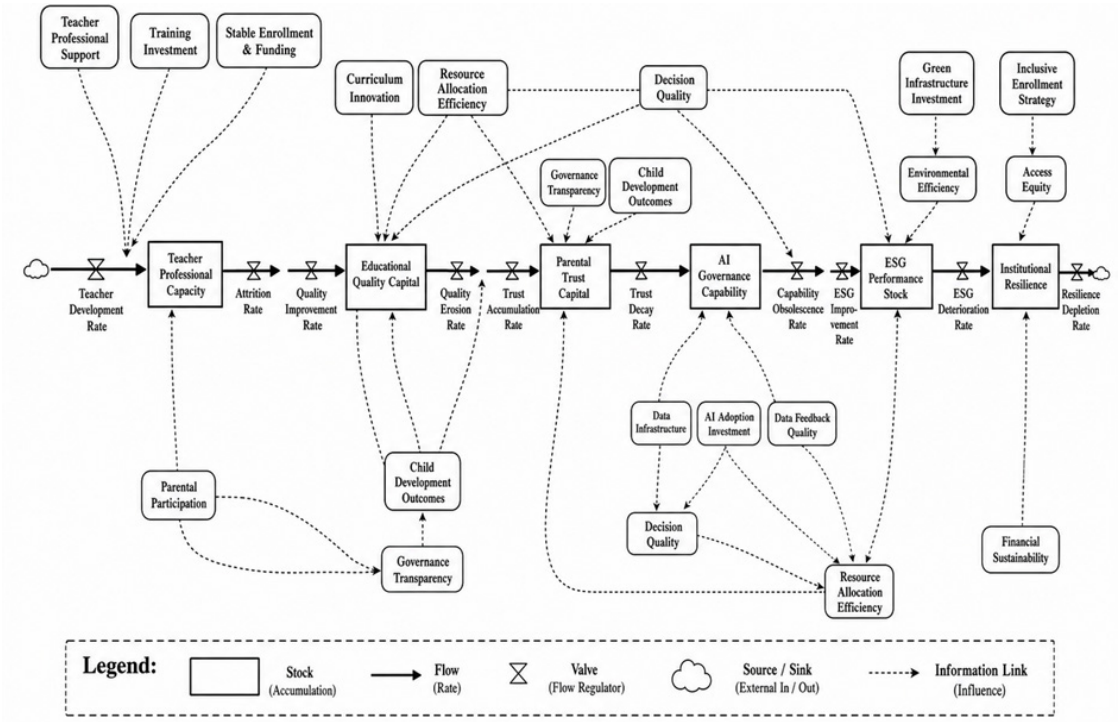


Figure 3. Stock-and-Flow Model of Smart and Sustainable Preschool Governance.

4.4 Scenario-based Analysis Results

As shown in Table 4, the scenario-based analysis suggests that under the S0 baseline scenario, education quality and parental trust only accumulate slowly, and are easily affected by teacher workload and financial constraints. S1 green campus investments can improve environmental performance, but short-term cost pressures are higher. S2 inclusive enrollment can improve educational equity, but the lack of subsidies and teacher allocation may lead to quality dilution. S3 transparent budget governance and parent communication can quickly increase trust capital. The S4 scenario "AI-Powered Integrated ESG Governance" is expected to generate the strongest long-term governance effects because it simultaneously enhances data analysis, policy forecasting, resource allocation, and performance tracking. Zawacki-Richter et al. (2019) pointed out that AI education applications have covered prediction, assessment, and support services, but they still need to incorporate educators' roles and governance ethics. Therefore, this study positions AI as a human-centered decision support rather than an automated alternative.

Table 4. Scenario-based Analysis Results of AI-enabled ESG Preschool Governance

Scenario	Scenario name	Result
S0	Baseline	Without additional policy intervention, various indicators only showed slow growth, and the overall improvement was limited.
S1	Green Infrastructure Investment	S1 is expected to primarily improve environmental and ESG-related performance, and it can moderately improve the quality of education and organizational resilience.
S2	Inclusive Enrollment Strategy	It helps improve educational equity and overall performance, but the improvement is relatively limited.
S3	Transparent Governance	The effect of improving trust in parents is the most significant, and simultaneously driving education quality and governance performance.
S4	Integrated AI-enabled ESG Governance	The integrated scenario is expected to improve all four dimensions more comprehensively than single-policy scenarios, indicating that integrating AI, ESG, and governance mechanisms can produce the strongest long-term synergies.

Summary of key findings

Overall, this study has five main findings (Table 5): First, ESG can be translated into environmental, social, and governance indicators for preschool governance; second, AI can enhance data-driven decision-making and governance knowledge transformation in preschools; third, teacher professional capacity and parental trust are core stocks of smart and sustainable preschool governance; Fourth, there is a policy trade-off between sustainable investment, educational equity and financial pressures; Fifth, a governance model that integrates AI, ESG, and system dynamics can enhance the long-term resilience and accountability of preschools more effectively than a single policy tool. Therefore, system dynamics is not only a model tool, but also a knowledge method for preschool administrators to conduct policy learning and strategy simulation (Groff, 2013; Pedamallu et al., 2010).

Table 5. Summary of Key Findings

Finding	Description
ESG can be translated into preschool governance indicators	Environmental, social, and governance goals can be operationalized in preschool management.
AI can enhance governance knowledge transformation	Administrative, learning, and stakeholder data can be converted into decision-support knowledge.
Teacher professional capacity and parental trust are core stocks	These two accumulated resources shape long-term preschool quality and legitimacy.
Sustainability investment involves policy trade-offs	Green investment, educational equity, and financial pressure interact dynamically.
Integrated AI–ESG–SD governance performs best	The integrated scenario produces stronger resilience and accountability than single-policy interventions.

5 Discussion

5.1 Theoretical contributions

The main theoretical contribution of this study is to extend ESG principles from corporate sustainability governance to preschool education governance, and to define preschools as early learning organizations that combine educational quality, social equity, and sustainability responsibilities. Existing research on early childhood sustainable education emphasizes children's participation, life practice, and environmental learning, pointing out that early childhood education is an important field for promoting sustainable social transformation. This study further elevates this perspective to the level of organizational governance, demonstrating that ESG can be used as an analytical framework for environmental management, teacher support, parental trust, and governance transparency in preschools (Engdahl & Furu, 2022).

5.2 Practical implications

In practice, this study suggests that preschool principals and education authorities can track key variables such as energy use, educational equity, teacher workload, parent engagement, and budget transparency through ESG indicator dashboards. System dynamics can help preschool leaders identify long-term feedback relationships between education quality, teacher pressure, resource allocation, and parental trust, avoiding

decision-making based solely on short-term enrollment or financial metrics. Groff (2013) pointed out that system dynamics models can reveal hidden dynamics in education policies and support future scenario design, making them suitable for preschool ESG policy simulations.

5.3 The significance of AI and ESG integration

The integration of AI and ESG is significant in transforming scattered administrative, instructional, resource-related, and parental feedback data in preschools into actionable governance knowledge. Alavi and Leidner believe that knowledge management systems can support the creation, storage, transfer, and application of organizational knowledge; AI can therefore serve as a knowledge base for ESG governance in preschools, assisting in predicting needs, allocating resources, and conducting scenario simulations (Alavi & Leidner, 2001). However, AI education applications still need to focus on educator roles, data ethics, and governance accountability, and avoid simplifying early childhood education to algorithmic management (Zawacki-Richter et al., 2019). Based on this, this study responds to the AI-driven analytics and ESG-aligned governance framework proposed in the previous summary, emphasizing that AI should be used as a human-centered decision support tool, rather than replacing the professional judgment of teachers and principals.

6 Conclusion

6.1 Research conclusions

This study constructs a preschool intelligent sustainable governance model of "ESG Principles-AI Decision Support-System Dynamic Simulation", pointing out that preschools are not only care and educational settings but also early learning organizations that cultivate sustainable literacy and social resilience. Research shows that ESG can translate into core variables such as environmental resource efficiency, educational equity, teacher support, parental trust, and governance transparency. AI can support enrollment forecasting, resource allocation, risk warning, and policy simulation, forming a data-driven governance knowledge system. The system dynamics model further reveals the long-term feedback relationship between teacher quality, child development, parental trust, financial resources, and sustainable investment, echoing the AI-driven analytics and ESG-aligned governance framework highlighted in the previous summary. Early childhood education research also points out that early education is an important field for promoting sustainable social transformation. System dynamics can assist educational administrators in analyzing hidden structures and future scenarios (Engdahl & Furu, 2022; Groff, 2013).

Overall, the study contributes an integrative governance perspective in which ESG defines the normative goals, AI enables data-driven knowledge transformation, and system dynamics supports policy learning under complexity.

6.2 Study limitations

This study is a conceptual modeling and scenario simulation study, and has not yet been used for parameter calibration and model verification based on long-term data from actual preschools. Secondly, differences in early childhood education systems, subsidy structures, teacher conditions, and ESG norms in different countries or regions may affect the extrapolation of models. Third, the introduction of AI into preschool governance involves data privacy, algorithm bias, educator roles, and accountability mechanisms, and more rigorous ethics and governance designs must be incorporated in the future. AI education research has emphasized that the role of educators and ethical risks must be paid attention to when introducing AI in educational fields, rather than just pursuing technological efficiency (Zawacki-Richter et al., 2019).

6.3 Future research directions

Future research can collect administrative, teacher, and parental feedback and sustainability performance data from multiple public and private preschools for model calibration, sensitivity analysis, and cross-case comparison. Secondly, an AI-enabled ESG dashboard can be developed to assist principals in real-time monitoring of education quality, teacher workload, resource efficiency, and parental trust. Finally, in the future, the perspectives of principals, teachers, parents, and competent authorities can also be included to examine the usability, trustworthiness, and ethical appropriateness of smart governance models in practical decision-making.

7 Declaration of Ethics

This study is based on conceptual modeling and scenario analysis. No human participants, children's personal data, or identifiable institutional records were collected. Therefore, formal ethical approval was not required at this stage. Future empirical validation involving preschool stakeholders will follow institutional ethical review procedures and data protection requirements.

8 AI Declaration

Generative AI tools were used to support language refinement, structural organization, and preliminary literature organization. The authors reviewed and verified all arguments, citations, interpretations, and final content, and take full responsibility for the accuracy and integrity of the manuscript.

References

Alavi, M., & Leidner, D. E. (2001). Knowledge management and knowledge management systems: Conceptual foundations and research issues. MIS Quarterly, 25(1), 107-136. https://doi.org/10.2307/3250961

Alfredo, R., Echeverria, V., Jin, Y., Yan, L., Swiecki, Z., Gašević, D., & Martinez-Maldonado, R. (2024). Human-centred learning analytics and AI in education: A systematic literature review. *Computers and Education: Artificial Intelligence*, *6*, 100215. https://doi.org/10.1016/j.caeai.2024.100215

Borg, F., & Samuelsson, I. P. (2022). Preschool children's agency in education for sustainability: The case of Sweden. E*uropean Early Childhood Education Research Journal*, *30*(1), 147-163. https://doi.org/10.1080/1350293X.2022.2026439

Brunsek, A., Perlman, M., McMullen, E., Falenchuk, O., Fletcher, B., Nocita, G., ... & Shah, P. S. (2020). A meta-analysis and systematic review of the associations between professional development of early childhood educators and children's outcomes. *Early Childhood Research Quarterly*, *53*, 217-248. https://doi.org/10.1016/j.ecresq.2020.03.003

Engdahl, I., & Furu, A. C. (2022). Early childhood education: A vibrant arena in the complex transformation of society towards sustainability. *International Journal of Early Childhood*, *54*(1), 1-12. https://doi.org/10.1007/s13158-022-00323-0

Finatto, C. P., Fuchs, P. G., Aguiar Dutra, A. R., & Guerra, J. B. S. D. A. (2024). Environmental, social, governance and sustainable development goals: promoting sustainability in universities. I*nternational Journal of Sustainability in Higher Education*, *25*(6), 1121-1136. https://doi.org/10.1108/IJSHE-11-2022-0361

Groff, J. S. (2013). Dynamic systems modeling in educational system design & policy. *Journal of New Approaches in Educational Research*, *2*(2), 72-81. https://doi.org/10.7821/naer.2.2.72-81

Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science*, *312*(5782), 1900-1902. https://doi.org/10.1126/science.1128898

Kennedy, M. (2011, July). A review of system dynamics models of educational policy issues. *Proceedings of the International Conference of the System Dynamics Society*, *Washington*, *DC*, *USA.*

Lim, C.-S., & Kim, Y.-E. (2025). The impact of ESG management in daycare Centers on childcare teachers' awareness of social Issues and environmentally friendly educational attitudes. *Journal of Child-Centered Practice*, *20*(2), 69–82. https://doi.org/10.53526/KACP.2025.20.2.69

Nores, M., & Barnett, W. S. (2010). Benefits of early childhood interventions across the world:(Under) Investing in the very young. *Economics of Education Review*, *29*(2), 271-282. https://doi.org/10.1016/j.econedurev.2009.09.001

Otero-Mayer, A., González-Benito, A., Gutiérrez-de-Rozas, B., & Expósito-Casas, E. (2025). Family involvement in early childhood education: a systematic review of its measurement. *Early Childhood Education Journal*, *1*-22. https://doi.org/10.1007/s10643-025-02024-4

Pedamallu, C., Ozdamar, L., Ganesh, L., Weber, G. W., & Kropat, E. (2010). A system dynamics model for improving primary education enrollment in a developing country. Organizacija, 43(3), 90-101. https://doi.org/10.2478/v10051-010-0010-5

Rezaei, M. (2025). Artificial intelligence in knowledge management: Identifying and addressing the key implementation challenges. *Technological Forecasting and Social Change*, *217*, 124183. https://doi.org/10.1016/j.techfore.2025.124183

Rosli, M. H., & Azmi, N. A. (2025). Integrating Environmental, Social and Governance (ESG) into Higher Education: Voices of Accounting Undergraduates. *Information Management and Business Review*, *17*(3(I)S), 127-135. https://doi.org/10.22610/imbr.v17i3(I)S.4677

Sterman, J. D. (2000). *Business dynamics: Systems thinking and modeling for a complex world*. Irwin/McGraw-Hill.

Strielkowski, W., Grebennikova, V., Lisovskiy, A., Rakhimova, G., & Vasileva, T. (2025). AI-driven adaptive learning for sustainable educational transformation. *Sustainable Development*, *33*(2), 1921-1947. https://doi.org/10.1002/sd.3221

Vodopivec, J. L., & Šindic, A. (2025). Early Childhood Education for Sustainability in Preschools with Eco-Programs: Implementation and Learning Outcomes. *Journal of Baltic Science Education*, *24*(1), 133-148. https://doi.org/10.33225/jbse/25.24.133

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education–where are the educators?. *International Journal of Educational Technology in Higher Education*, *16*(1), 39. https://doi.org/10.1186/s41239-019-0171-0

How Does Knowledge Externalization Quality Shape the ESG–Financial Performance Nexus: An Explainable AI Analysis under Mandatory Disclosure Regime

Bin-Wei Hsu[1], Hsiu-Yi Lin[2]
[1] Department of Industrial Management, National Pingtung University of Science and Technology
[2] Department of Finance, National Pingtung University

Abstract

How organizations transform tacit sustainability knowledge into externally interpretable explicit information — and whether the quality of this transformation process influences capital market valuation — represents a core question that knowledge management and capital market research have yet to fully integrate. This study repositions corporate ESG (Environmental, Social, and Governance) disclosure as an institutional output of organizational knowledge externalization, further conceptualizing it as knowledge-based signals that enter market valuation mechanisms through investor interpretation. Grounded in Nonaka and Takeuchi's (1995) knowledge externalization theory and signaling theory, this study constructs a Knowledge Externalization Quality (KE Quality) indicator to measure the extent to which firms convert tacit sustainability knowledge into verifiable and comparable explicit information. Using Taiwan's mandatory ESG disclosure policy as an institutional shock, this study employs the Taiwan Economic Journal TESG database combined with financial and market data to build a longitudinal panel of 1,033 listed companies with 8,672 observations, examining how KE Quality shapes the market pricing of ESG signals. The study adopts fixed-effects OLS regression with robust standard errors, and triangulates findings through lagged specifications, Difference-in-Differences (DID) design, and XGBoost explainable AI (SHAP) analysis.

Empirical results reveal four key findings. First, KE Quality among Taiwanese firms improved significantly from 0.481 in 2016 to 0.683 in 2024, while ESG substantive scores grew only marginally (approximately 3.7%), indicating a pervasive "formal externalization" phenomenon under institutional pressure. Second, KE Quality exerts a significant positive moderating effect on market valuation (Tobin's Q; $\beta=0.033$, $p<0.01$), but has no significant impact on short-term stock returns, reflecting a temporally delayed market interpretation of ESG knowledge signals. Third, the signaling effects of Environmental (E) and Social (S) dimensions exhibit asymmetry: environmental disclosure enhances market premiums, while social disclosure compresses short-term accounting performance. Fourth, DID analysis shows that under mandatory disclosure, large firms experienced significant declines in both KE Quality (DID=−0.110, $p<0.01$) and market valuation (DID=−0.203, $p<0.01$), supporting the signal dilution hypothesis. Finally, the XGBoost model (test $R^2=0.615$) and SHAP analysis corroborate the regression findings, providing methodological triangulation. The theoretical contribution of this study lies in demonstrating that the economic value of ESG information does not derive from disclosure per se, but from the quality of its knowledge externalization signal and how institutional environments shape the interpretability and market pricing of that signal.

Keywords: Knowledge externalization; ESG disclosure; Market valuation; Signal theory; Mandatory disclosure

Paper type: Research paper

1 Introduction

When an organization's sustainability practices are codified into an ESG report, can the quality of this knowledge artifact — its completeness, verifiability, and the motivation behind its disclosure — be effectively interpreted by capital markets and reflected in firm valuation? This is among the most compelling questions that knowledge management theory can address in contemporary sustainability governance. This study positions ESG sustainability reports as institutional outputs of organizational knowledge externalization (Nonaka and Takeuchi, 1995), rather than mere information disclosure acts. Taiwan's Financial Supervisory Commission (FSC) has progressively mandated sustainability reporting since 2014, substantially expanding the mandatory scope after 2023, creating a natural experiment to observe how institutional intervention transforms the market signaling effects of knowledge externalization behavior. Over nine years, the sustainability report filing rate among Taiwan's listed firms surged from 33% to nearly 69%, yet substantive ESG score improvements lagged far behind — this gap represents the concrete manifestation of a disconnect between knowledge externalization quality and knowledge content. This study defines this as the "formal externalization" phenomenon and proceeds from this core problem.

Existing ESG research has concentrated primarily on the direct association between ESG scores and financial performance (Friede et al., 2015; Chen and Xie, 2022), or on mandatory disclosure's effects on capital markets (Krueger et al., 2024). These studies, however, treat ESG disclosure as a finished information product,

overlooking the production process itself — how organizations transform tacit sustainability knowledge into externally comparable explicit reports — which is the critical determinant of signal quality. This study's position is that ESG disclosure should be understood not as a reporting practice, but as a process of organizational knowledge externalization, where the quality of codification determines how knowledge is interpreted and priced in capital markets. Nonaka and Takeuchi (1995) established that the quality of the process by which organizations convert tacit knowledge into explicit knowledge determines the effectiveness of knowledge transfer. ESG sustainability reports are the concrete realization of this process — quality differences (voluntary vs. mandated disclosure, presence of third-party assurance, completeness of disclosure) constitute observable dimensions of knowledge codification effectiveness that should systematically influence external stakeholder valuations.

This study addresses three research questions: What is the trajectory of KE Quality evolution among Taiwanese firms during the ESG policy promotion period (2016–2024)? How does KE Quality moderate the relationship between ESG scores and financial performance, and does this moderation vary by ESG dimension and disclosure motivation? Does the financial effect of knowledge externalization behavior undergo structural change following mandatory disclosure intervention? The core theoretical contribution of this study to knowledge management literature lies in demonstrating that the market value of organizational knowledge is shaped not solely by knowledge content, but also by the quality of the knowledge externalization process under institutional constraints — a proposition that challenges the linear assumption that "more disclosure leads to higher market valuation" and establishes new boundary conditions for applying knowledge management theory in capital market contexts.

2 Literature Review

2.1 Knowledge Externalization and ESG Disclosure

Nonaka and Takeuchi (1995) established in their SECI (Socialization-Externalization-Combination-Internalization) knowledge spiral model that knowledge externalization — the conversion of tacit organizational knowledge into transmissible explicit knowledge — is a critical process whose quality directly determines the effectiveness and interpretability of knowledge transfer. In the ESG context, corporate sustainability practices (such as carbon management, labor relations, and governance mechanisms) are largely embedded in daily operations, characterized by high tacitness; the compilation of sustainability reports constitutes the institutionalized mechanism for systematically translating such tacit knowledge into externally observable information.

Established disclosure research and international reporting standards (e.g., the Global Reporting Initiative) indicate that high-quality disclosure typically exhibits completeness, comparability, and credibility (e.g., third-party assurance) — features that collectively form the basis of information interpretability and market trust. From a knowledge management perspective, these dimensions can be understood as concrete manifestations of knowledge externalization quality, reflecting the effectiveness of knowledge codification and transmission.

Signaling theory (Spence, 1973) provides a complementary framework: firms convey their internal capabilities and long-term orientation through ESG disclosure, and high-quality disclosure helps reduce information asymmetry, enabling investors to distinguish genuinely sustainable firms from those engaging in superficial compliance. In this sense, ESG disclosure can be understood as externalized knowledge signals through which investors conduct quality assessments and market valuation.

However, when disclosure transitions from voluntary to mandated, the economic implications of signals may undergo structural change. Christensen et al. (2021) indicate that mandatory disclosure regimes may erode the function of information as a differentiating signal, leading to signal dilution. In this context, markets no longer rely on "whether to disclose" as a basis for judgment, but shift toward quality differences in disclosure — making KE Quality the critical mechanism influencing market valuation. This study uses Taiwan's mandatory disclosure policy as an institutional shock and employs a Difference-in-Differences (DID) design to empirically test whether signal dilution effects exist in capital markets.

2.2 Mandatory Disclosure, ESG Quality, and Financial Performance

Existing research on ESG disclosure and corporate economic outcomes exhibits substantial heterogeneity. Friede et al. (2015), integrating more than 2,000 empirical studies, find that approximately 90% show a non-negative relationship between ESG and corporate financial performance, with most being positive, though effect sizes vary considerably across regions, industries, and measurement approaches. This result suggests that ESG

information is not a stable source of value; its economic effect depends on how markets interpret that information.

Further literature indicates that ESG information's impact on capital markets is primarily manifested in market valuation and cost of capital. Dhaliwal et al. (2011) find that voluntary non-financial disclosure helps reduce equity capital costs; Bolton and Kacperczyk (2021) demonstrate that markets price environmental risk differentially across firms, indicating that different ESG dimensions carry distinct valuation mechanisms. Investor composition also affects the transmission efficiency of ESG signals (Chen and Xie, 2022), further amplifying the heterogeneity of market responses.

On the institutional dimension, mandatory disclosure policy effects are also non-linearly positive. Krueger et al. (2024), drawing on cross-country data, find that mandatory ESG disclosure policies influence market liquidity and information environments, but effects vary substantially across firms. Christensen et al. (2021) further indicate that when disclosure becomes a regulatory obligation, firms may gravitate toward minimum compliance rather than improving information quality, thereby weakening disclosure's function as an effective signal.

Taken together, existing literature implies a core proposition that has yet to be fully examined: the market valuation effects of ESG disclosure depend not only on disclosure content, but on the quality of the knowledge externalization process, and on how institutional environments shape the interpretability of signals. This study accordingly proposes that under mandatory disclosure regimes, interfirm variation shifts from "whether to disclose" toward "disclosure quality," further shaping the pricing mechanism of ESG signals in capital markets.

2.3 Hypotheses

- *H1: During Taiwan's ESG policy promotion period (2016–2024), KE Quality shows a year-on-year upward trend, but the improvement in substantive ESG scores is smaller than the increase in disclosure behavior (formal externalization).*
- *H2: KE Quality positively moderates the effect of ESG scores on long-term market valuation (Tobin's Q), but has no significant immediate effect on short-term stock returns (market signal lag).*
- *H3: KE Quality's moderating effect on the ESG–financial performance relationship exhibits dimensional heterogeneity, with the Environmental (E) and Social (S) dimensions showing different directions and magnitudes of moderation across performance measures.*
- *H4: Voluntary early adopters exhibit significantly positive KE Quality moderation effects, while passive late adopters (firms coerced by policy) show significantly weaker or negative effects, reflecting that disclosure motivation shapes the financial effects of knowledge externalization.*

3 Methodology

3.1 Data and Sample

This study employs the Taiwan Economic Journal (TEJ) TESG Sustainability Development Index database, combined with TEJ Finance financial data and stock return data, to construct a longitudinal dataset spanning 2016 to 2024. The TESG database is structured around the SASB (Sustainability Accounting Standards Board) international framework, drawing on over 20 public information sources. Its distinctive feature is the ability to generate quantitative ESG scores even for firms that have not published sustainability reports — using external information such as ISO certifications and Ministry of Labor employment data — providing a systematic basis for comparing firms at different disclosure stages. The final sample covers 1,033 listed companies with 8,672 annual observations. Financial variables are Winsorized at the 1st and 99th percentiles; stock returns are measured using the trailing 12-month return as of each December year-end.

3.2 Variable Definitions

Knowledge Externalization Quality (KE Quality) is a composite index constructed from three equally weighted (1/3 each) conditions, ranging from 0 to 1: (1) whether a sustainability report has been published (1=yes, 0=no), reflecting the firm's willingness to voluntarily externalize knowledge; (2) whether third-party assurance has been obtained (1=yes, 0=no), reflecting the verifiability of externalized knowledge; and (3) disclosure completeness (0 to 1), standardized from TESG event radar scoring, where higher scores indicate more complete disclosure and fewer negative events. These three dimensions collectively measure the extent to which organizations effectively codify (codification) and verify (verifiability) tacit sustainability knowledge — not merely the disclosure act itself. KE Quality reflects the quality of the knowledge conversion process, while ESG scores reflect the content level of knowledge; the two are conceptually independent, and this study employs interaction terms

to capture their joint effects. Robustness analyses under five alternative weighting schemes (Analysis 10) confirm the direction and significance of interaction coefficients throughout.

Primary dependent variables include ROA (Return on Assets), ROE (Return on Equity), and Tobin's Q; stock returns serve as an additional measure of immediate signal response. Independent variables include the TESG composite score and three sub-dimension scores — Environmental (E), Social (S), and Governance (G) — all standardized. Control variables include firm size (log of total assets), debt ratio, and firm age.

Disclosure motivation grouping: based on the year of first sustainability report publication, early voluntary adopters (adopted before 2023, N=4,539 observations) and late passive adopters (joined under policy pressure after 2023, N=1,209 observations) are defined. The DID design uses paid-in capital as the treatment threshold: firms with capital ≥ NTD 2 billion form the treatment group (subject to earlier mandatory requirements, N=608 observations), while firms below this threshold form the control group. The DID event corresponds to the 2023 expansion of mandatory disclosure policy.

3.3 Analytical Methods

The primary analysis employs OLS regression incorporating year and industry fixed effects with HC1 (Heteroscedasticity-Consistent) robust standard errors. The core model specification is: Financial Performance = $\alpha + \beta_1 \times$TESG Score + $\beta_2 \times$KE Quality + $\beta_3 \times$(TESG×KE Quality) + $\beta_4 \times$Controls + Fixed Effects + ε. Robustness checks include: lagged variables (t−1) to address endogeneity, E/S/G sub-dimension robustness tests, three alternative time windows, and sensitivity analysis under five KE Quality weighting schemes.

4 Results

4.1 Descriptive Statistics

Table 1 presents descriptive statistics for core variables. The mean TESG composite score is 56.07, with E, S, and G sub-dimension scores at 57.43, 57.01, and 54.46, respectively — the governance dimension systematically lower than environmental and social. The median KE Quality is 0.333, implying that more than half of Taiwan's listed companies during the study period possessed only the disclosure completeness dimension and had not yet published sustainability reports. The distribution of early voluntary adopters (N=4,539) and late passive adopters (N=1,209) provides sufficient statistical power for the subsequent motivation heterogeneity analysis.

Table 1: Descriptive Statistics of Core Variables (2016–2024)

Variable	Mean	SD	Min	Max	Median	N
TESG Composite Score	56.07	8.58	29.89	83.73	55.62	8,672
E Score (Environmental)	57.43	12.79	26.02	92.16	55.67	8,672
S Score (Social)	57.01	11.78	26.58	91.00	56.43	8,672
G Score (Governance)	54.46	10.15	21.85	81.54	54.79	8,672
KE Quality	0.563	0.281	0.000	1.000	0.333	8,672
Has Sustainability Report	0.473	0.499	0	1	0	8,672
Has Third-party Assurance	0.277	0.447	0	1	0	8,672
ROA (%)	4.26	7.44	−22.83	27.60	3.92	8,672
ROE (%)	7.63	14.03	−48.73	48.29	7.95	8,672
Tobin's Q	1.221	0.891	0.090	5.616	0.960	8,672
Firm Size (ln)	16.25	1.65	9.76	23.35	15.98	8,672
Debt Ratio (%)	44.83	19.95	0.48	99.76	44.55	8,672
Firm Age (years)	40.32	15.85	2	80	40	8,663

Source: This study. Financial variables Winsorized at 1st and 99th percentiles.

4.2 Trajectory of KE Quality Over Time

Figure 1 shows that KE Quality among Taiwanese firms rose continuously from 0.481 in 2016 to 0.683 in 2024, a cumulative improvement of 42.0% over nine years. The sustainability report filing rate climbed from 33.0% to 68.9%, and the third-party assurance rate from 16.5% to 44.2%, clearly demonstrating the policy's effect. However, the TESG composite score improved only from 55.36 to 57.42 over the same period — a mere 3.7% gain, far below the increase in disclosure behavior. The comparison between the pre-policy period (2016–2022) and the post-policy period (2023–2024) shows KE Quality rising by 0.133, while substantive ESG scores improved by only 1.761 points; the gap between these two rates of improvement is statistically significant ($p<0.01$), supporting H1's formal externalization proposition.

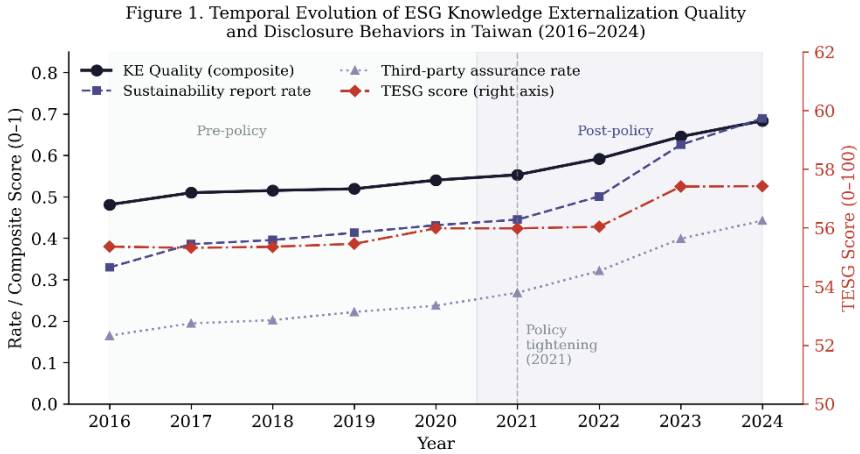


Figure 1. Annual Evolution of ESG Knowledge Externalization Quality and Disclosure Behavior among Taiwan Listed Firms (2016–2024). Left axis: KE Quality and disclosure rates (0–1); Right axis: TESG score (50–62).

4.3 Regression Results

Table 2 presents results from four regression models (DV=ROA). Model 1 shows TESG scores have a significant positive effect on ROA (β=0.616, p<0.01). Model 2, adding KE Quality, shows KE Quality itself exerts a significant negative effect on ROA (β=−0.242, p<0.05), reflecting the short-term sustainability costs borne by proactive disclosers. The core interaction term in Model 3 (TESG×KE Quality) is −0.161 (p<0.1), indicating weakly significant negative moderation — in firms with high disclosure quality, the short-term financial effect of ESG scores is partially offset. Model 4 results (moved to robustness appendix; see Section 4.6) present more pronounced policy moderation: the post-policy × KE Quality interaction is significantly negative (β=−0.397, p<0.01), representing structural weakening of KE Quality's effect on ROA after policy intervention, consistent with the formal externalization hypothesis.

Table 2: Panel Regression Results (DV=ROA)

Variable	Model 1 Baseline	Model 2 +KE Quality	Model 3 Interaction	Model 4 Policy Mod.
TESG Score β	0.616***	0.709***	0.716***	0.741***
KE Quality β	—	−0.242**	−0.216**	−0.111
Interaction (TESG×KE)	—	—	−0.161*	—
Post-policy Dummy	—	—	—	1.085***
Post-policy × KE Quality	—	—	—	−0.397***
R^2	0.177	0.178	0.178	0.168
N	8,663	8,663	8,663	8,663

****p<0.01 **p<0.05 *p<0.1. All models include year and industry fixed effects with HC1 robust standard errors. Source: This study.*

Table 3 presents the core model comparison across four dependent variables. KE Quality's moderating effect on Tobin's Q is significantly positive (β=0.033, p<0.01) and remains robust in the lagged variable model (β=0.022, p<0.05), ruling out reverse causality concerns and supporting the first part of H2. By contrast, the interaction term for short-term stock returns is not significant (β=−0.436), indicating market lag in responding to ESG disclosure quality — supporting the second part of H2.

Table 3: Core Interaction Model — Comparison Across Four Dependent Variables

DV	TESG Score β	KE Quality β	Interaction β	R^2	N
ROA	0.716***	−0.216**	−0.161*	0.178	8,663
ROE	1.332***	−0.420**	−0.194	0.145	8,663
Tobin's Q	0.068***	0.014	0.033***	0.259	8,663
Stock Return	−0.871	−0.383	−0.436	0.134	8,663

***p<0.01 **p<0.05 *p<0.1. All models include year and industry fixed effects. Source: This study.

4.4 E/S/G Dimensional Heterogeneity

Table 4 presents results substituting E, S, and G sub-dimension scores for the TESG composite score, with Figure 2 displaying interaction coefficients by dimension. The Environmental (E) dimension interaction is significantly positive on Tobin's Q (β=0.050, p<0.01), indicating that transparent disclosure of environmental knowledge yields a market premium. Conversely, the Social (S) dimension interaction is significantly negative on both ROA (β=−0.327, p<0.01) and ROE (β=−0.611, p<0.01), suggesting that transparent social responsibility disclosure compresses short-term accounting performance. The Governance (G) dimension is also significantly positive on Tobin's Q (β=0.025, p<0.01). This E/S/G asymmetry in moderation directions supports H3, demonstrating that the financial effects of KE Quality are highly dependent on the type of sustainability knowledge being externalized.

Table 4: KE Quality Moderation Effects by E/S/G Sub-dimension

Dim.	DV	ESG Score β	KE Quality β	Interaction β	R^2	N
E	ROA	0.468***	−0.076	0.023	0.175	8,663
E	ROE	0.867***	−0.166	0.222	0.143	8,663
E	Tobin's Q	0.013	0.041***	0.050***	0.258	8,663
S	ROA	0.171*	0.081	−0.327***	0.174	8,663
S	ROE	0.261	0.178	−0.611***	0.141	8,663
S	Tobin's Q	0.025**	0.043***	0.005	0.256	8,663
G	ROA	0.537***	−0.052	−0.029	0.177	8,663
G	ROE	0.986***	−0.096	−0.015	0.144	8,663
G	Tobin's Q	0.072***	0.028**	0.025***	0.261	8,663

***p<0.01 **p<0.05 *p<0.1. All models include year and industry fixed effects. Source: This study.

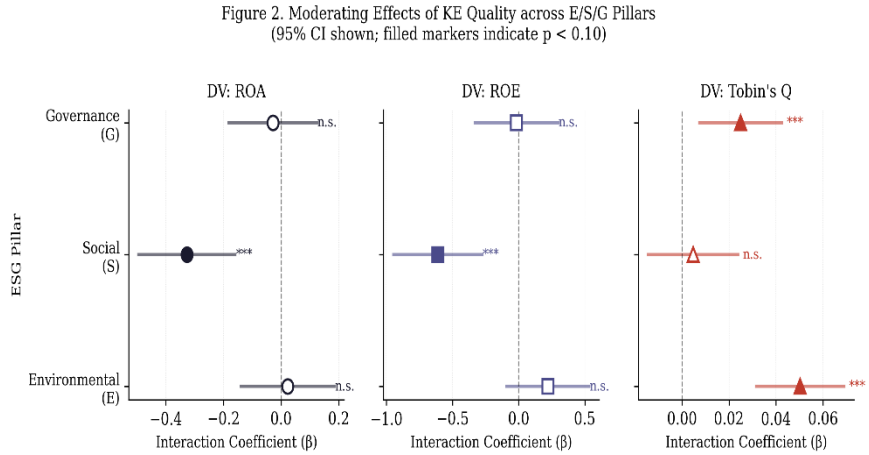


Figure 2. Coefficient Plot of KE Quality Moderation Effects by E/S/G Dimension (with 95% Confidence Intervals).

4.5 Motivation Heterogeneity and DID Analysis

Table 5 Panel A presents motivation heterogeneity results. Early voluntary adopters show a significantly positive interaction (β=0.061, p<0.01), while late passive adopters show an interaction near zero (β=−0.009, n.s.), with KE Quality itself exhibiting a significant negative effect on Tobin's Q (β=−0.172, p<0.05). This contrast directly supports H4: disclosure motivation determines the market effects of KE Quality — voluntary disclosure functions as a signal of genuine sustainability capability with positive moderating power, whereas passive disclosure does not carry this signaling function in the eyes of the market.

Table 5 Panel B presents DID results. Using the NTD 2 billion capital threshold, treatment group firms experienced a significant decline in KE Quality (DID=−0.110, p<0.01) and a corresponding drop in Tobin's Q (DID=−0.203, p<0.01) following the 2023 policy intervention. This result appears to contradict the main regression finding that KE Quality positively moderates ESG's effect on Tobin's Q; the theoretical interpretation is addressed in Section 5.

Table 5: Motivation Heterogeneity (Panel A) and DID Analysis (Panel B)

Analysis	TESG β	KE Quality β	Interaction β	R^2 / N
Panel A: Motivation Heterogeneity (DV=Tobin's Q)				
Early voluntary (N=4,539)	0.017	0.080***	0.061***	R^2=0.312
Late passive (N=1,209)	0.047	−0.172**	−0.009	R^2=0.285
Panel B: DID Analysis (Capital Threshold ≥ NTD 2 billion)				
DV=KE Quality	−0.036***(treat)	0.089***(post)	−0.110***(DID)	R^2=0.468 / N=8,663
DV=ROA	−4.434***(treat)	1.019***(post)	−1.153**(DID)	R^2=0.175 / N=8,663
DV=Tobin's Q	0.094**(treat)	0.175***(post)	−0.203***(DID)	R^2=0.247 / N=8,672

***p<0.01 **p<0.05 *p<0.1. Panel A DV=Tobin's Q, industry fixed effects included. Panel B DID design: treat=capital≥NTD 2B, post=post-2023, DID=treat×post interaction. Source: This study.

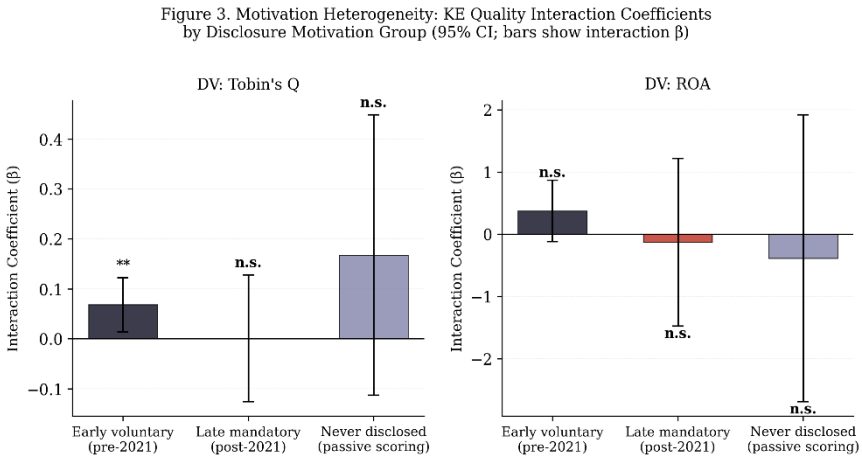


Figure 3. Motivation Heterogeneity: KE Quality Moderation Effects for Early Voluntary vs. Late Passive Adopters.

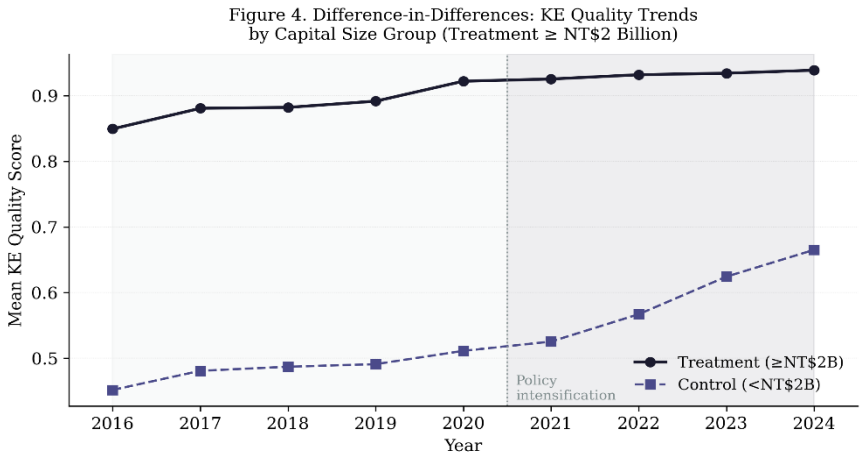


Figure 4. Difference-in-Differences Analysis: Annual KE Quality Trends by Capital Size Group (2016–2024).

4.6 Robustness Checks

This study conducts five categories of robustness checks, all supporting the main findings. First, in the lagged variable model (t−1), the Tobin's Q interaction remains significant (β=0.022, p<0.05), ruling out reverse causality. Second, across three alternative time windows (2018–2024, 2019–2024, and 2016–2022), interaction coefficients remain significant (β=0.029–0.042, p<0.01), confirming results are not period-specific. Third, all five alternative KE Quality weighting schemes yield significant interactions (β=0.031–0.036, p<0.01), demonstrating robustness of index construction. Fourth, industry heterogeneity analysis shows the moderation effect is most pronounced in chemicals/biomedical (β=0.094, p<0.1) and construction materials (β=0.094, p<0.05) sectors. Model 4 (policy moderation) is relegated to the appendix as a robustness reference due to reduced statistical

power following the post-2023 sample restriction (2,051 observations; policy_ke interaction: β=−0.263, p=0.134); causal inference for policy effects is carried by the DID design in Analysis 13.

4.7 Explainable AI Diagnostics: XGBoost and SHAP Analysis

To supplement the limitations of linear regression models and provide a nonlinear perspective, this study employs an XGBoost machine learning model to predict KE Quality, and applies SHAP (SHapley Additive exPlanations) explainable AI to diagnose the marginal contributions of each predictor. Input features include the TESG composite score, E/S/G sub-scores, firm size, debt ratio, firm age, ROA, ROE, Tobin's Q, post-policy dummy, and paid-in capital — 12 variables in total. Data are split 80/20 by year order (training: 6,930; test: 1,733) to prevent future information leakage. It is worth noting that conceptual independence does not imply empirical uncorrelatedness — firms with higher ESG content levels may also possess greater capacity for high-quality externalization; however, at the theoretical level, knowledge content (ESG scores) and knowledge externalization process quality (KE Quality) remain distinct constructs, and this distinction is a core contribution of this study.

The XGBoost model achieves R^2=0.615 and RMSE=0.183 on the test set, indicating considerable predictive power. Figure 6 presents the SHAP results: Panel A shows the feature importance bar chart, and Panel B shows the SHAP distribution plot for the top six features (red=high feature value, blue=low). SHAP analysis identifies the TESG composite score (mean SHAP=0.070) and post-policy dummy (mean SHAP=0.068) as the two most important predictors of KE Quality, with nearly equal contributions — indicating that ESG capability and institutional pressure exert roughly comparable driving forces on knowledge externalization behavior. Firm size ranks third (mean SHAP=0.063), reflecting the systematic advantage of large firms in disclosure resources and capacity.

Notably, the SHAP values for the post-policy dummy are almost entirely positive in Panel B, but taken together with DID results, this pattern is consistent with the interpretation that mandatory policy raised KE Quality's surface values without fostering genuine improvements in organizational knowledge externalization capacity — consistent with the predictions of this study's formal externalization hypothesis. Furthermore, E scores and S scores both show positive SHAP directions, but E score influence is more concentrated than S score influence, consistent with the regression finding that the environmental dimension moderation (β=0.050, p<0.01) outperforms the social dimension (β=0.005, n.s.), providing methodological triangulation. Financial performance indicators (ROA, ROE, Tobin's Q) all have mean SHAP values below 0.007, suggesting KE Quality is primarily driven by ESG characteristics and institutional environment, with limited direct association with short-term financial performance.

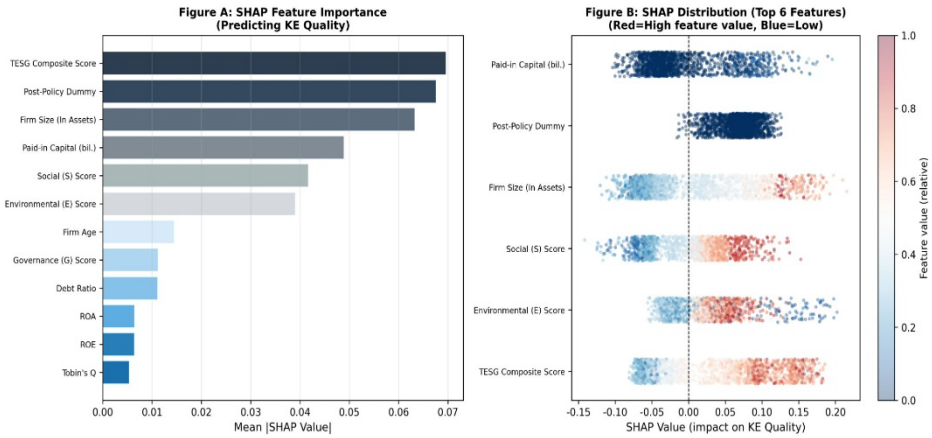


Figure 6. SHAP Explainable AI Analysis: Feature Importance (Panel A) and Distribution Plot (Panel B) for XGBoost Prediction of KE Quality.

5 Discussion

This study's most striking finding is the apparent contradiction between main regression and DID results: main regressions show that higher KE Quality amplifies ESG scores' positive effect on Tobin's Q, yet DID results show that firms subject to mandatory disclosure — large-capitalization firms — experienced simultaneous declines in KE Quality (DID=−0.110, p<0.01) and Tobin's Q (DID=−0.203, p<0.01) following the 2023 policy intervention. This contradiction points toward the signal dilution mechanism. Christensen et al. (2021) indicate

that mandatory disclosure policies may generate unintended consequences — when disclosure becomes a regulatory obligation rather than a voluntary act of knowledge sharing, markets systematically discount ESG signals from mandated disclosers, as these signals lose their differentiating function. This study's DID results directly echo this argument: among firms covered by mandatory policy, even as KE Quality values continued to improve, market valuations declined, reflecting investors' qualitative distinction between coerced and voluntary disclosure. Regarding causal inference, the DID design captures variation in regulatory exposure, but the capital threshold may simultaneously reflect firm size heterogeneity; accordingly, these results should be interpreted as policy-associated findings rather than strict causal identification. Robustness analyses with multiple thresholds (≥NTD 500M, ≥NTD 1B, ≥NTD 2B) partially mitigate this concern.

Motivation heterogeneity findings further reinforce the above interpretation. Early voluntary adopters show significantly positive KE Quality moderation ($\beta=0.061$, $p<0.01$), while the effect for late passive adopters nearly disappears ($\beta=-0.009$, n.s.), with KE Quality itself negatively associated with market valuation ($\beta=-0.172$, $p<0.05$). From the perspective of Nonaka and Takeuchi's (1995) knowledge externalization theory, voluntary disclosers' externalization behavior stems from the organization's intrinsic recognition of knowledge value, allowing externalization quality to authentically reflect sustainability capability; passive disclosers' externalization occurs under external institutional pressure, with a looser connection between quality and substantive capability, making it difficult for the market to extract useful signals.

The signal lag finding reveals a distinctive feature of Taiwan's capital market. The interaction for short-term stock returns is insignificant, while Tobin's Q interaction is significant in both full-sample and lagged variable models, indicating that market responses to ESG disclosure quality are not immediate but materialize in long-term valuation over time. This result may be associated with the incomplete maturity of Taiwan's market mechanism for processing ESG knowledge quality signals; Chen and Xie (2022) provide more general cross-market evidence that investor composition influences ESG signal transmission efficiency, offering theoretical context for this study's finding. Figure 5 visually presents this short-term versus long-term market response differential.

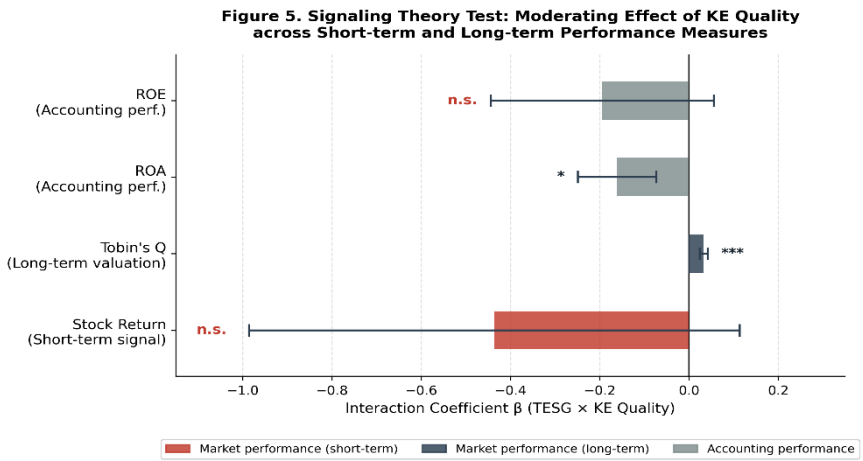


Figure 5. Signaling Theory Test: Moderating Effect of KE Quality across Short-term and Long-term Performance Measures.

6 Conclusion

This study repositions ESG sustainability disclosure as an institutional process of organizational knowledge externalization, employing Nonaka and Takeuchi's (1995) knowledge externalization theory as the core framework — combined with signaling theory and institutional theory — to systematically examine the evolution of KE Quality and its heterogeneous effects on capital market valuation during Taiwan's mandatory ESG disclosure policy period. The study's core knowledge management contributions are fourfold. First, it operationalizes KE Quality for the first time as a quantifiable composite measure of "knowledge codification effectiveness," explicitly demarcating the conceptual boundary between knowledge content (ESG scores) and knowledge externalization process quality (KE Quality), and empirically demonstrates that moderating effects vary significantly by ESG dimension (E vs. S), performance measurement dimension (market valuation vs. accounting performance vs. immediate market response), and disclosure motivation (voluntary vs. passive) — establishing specific boundary conditions for applying knowledge management theory in ESG contexts. Second, the DID design provides causal evidence for "institutional intervention distorting knowledge signals" — mandatory disclosers' KE Quality (DID=−0.110, $p<0.01$) and market valuation (DID=−0.203, $p<0.01$) both declined

following the 2023 policy, directly challenging the mainstream assumption that "more disclosure necessarily generates higher knowledge value," and providing a knowledge management interpretive framework for signal dilution theory. Third, the XGBoost explainable AI model (test R^2=0.615) and SHAP analysis constitute a diagnostic toolkit for knowledge externalization behavior, with TESG scores and policy dummy as the most important predictors — corroborating regression findings and demonstrating the methodological potential of explainable AI in knowledge management research. Fourth, the delayed market response finding indicates that Taiwan's capital market has yet to fully mature in its mechanism for interpreting ESG knowledge quality signals; the knowledge value realization pathway is primarily through long-term valuation, with short-term stock prices not yet fully reflecting differences in knowledge externalization quality.

For policymakers, this study recommends that mandatory disclosure policy be supplemented with quality incentive mechanisms rather than merely requiring formal compliance; a regulatory framework that differentiates between voluntary and passive disclosers would help preserve the differentiating value of ESG signals. For investors, when assessing corporate ESG performance, the motivational character of disclosure should be distinguished, with priority attention given to third-party assured environmental knowledge disclosure. Primary limitations include: the sample being restricted to Taiwan's listed companies, the three dimensions of KE Quality being equally weighted, and the use of static capital levels as DID grouping criteria — areas for future research to deepen.

7 Ethics Statement

This study uses publicly available commercial databases from the Taiwan Economic Journal (TEJ) and public financial disclosures from listed companies. It does not involve human subjects research and does not require Institutional Review Board approval.

8 AI Tool Usage Statement

This study used Claude (Anthropic) as an assistive tool during research design planning and initial drafting stages, for logical structuring of research design and text optimization. All statistical analyses were independently executed by the authors, and all interpretations of findings and academic judgments are the authors' responsibility.

References

Akerlof, G.A. (1970) "The Market for Lemons: Quality Uncertainty and the Market Mechanism", Quarterly Journal of Economics, Vol. 84, No. 3, pp. 488–500.

Bolton, P. and Kacperczyk, M. (2021) "Do Investors Care about Carbon Risk?", Journal of Financial Economics, Vol. 142, No. 2, pp. 517–549.

Chen, Z. and Xie, G. (2022) "ESG Disclosure and Financial Performance: Moderating Role of ESG Investors", International Review of Financial Analysis, Vol. 83, Article 102291.

Christensen, H.B., Hail, L. and Leuz, C. (2021) "Mandatory CSR and Sustainability Reporting: Economic Analysis and Literature Review", Review of Accounting Studies, Vol. 26, No. 3, pp. 1176–1248.

Dhaliwal, D.S., Li, O.Z., Tsang, A. and Yang, Y.G. (2011) "Voluntary Nonfinancial Disclosure and the Cost of Equity Capital: The Initiation of Corporate Social Responsibility Reporting", The Accounting Review, Vol. 86, No. 1, pp. 59–100.

DiMaggio, P.J. and Powell, W.W. (1983) "The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields", American Sociological Review, Vol. 48, No. 2, pp. 147–160.

Dumay, J., Bernardi, C., Guthrie, J. and Demartini, P. (2016) "Integrated Reporting: A Structured Literature Review", Accounting Forum, Vol. 40, No. 3, pp. 166–185.

Friede, G., Busch, T. and Bassen, A. (2015) "ESG and Financial Performance: Aggregated Evidence from More than 2000 Empirical Studies", Journal of Sustainable Finance and Investment, Vol. 5, No. 4, pp. 210–233.

Krueger, P., Sautner, Z., Tang, D.Y. and Zhong, R. (2024) "The Effects of Mandatory ESG Disclosure Around the World", Journal of Accounting Research, Vol. 62, No. 5, pp. 1795–1847.

Lins, K.V., Servaes, H. and Tamayo, A. (2017) "Social Capital, Trust, and Firm Performance: The Value of Corporate Social Responsibility during the Financial Crisis", Journal of Finance, Vol. 72, No. 4, pp. 1785–1824.

Nonaka, I. and Takeuchi, H. (1995) The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation, Oxford University Press, New York.

Nonaka, I., Toyama, R. and Konno, N. (2000) "SECI, Ba and Leadership: A Unified Model of Dynamic Knowledge Creation", Long Range Planning, Vol. 33, No. 1, pp. 5–34.

Spence, M. (1973) "Job Market Signaling", Quarterly Journal of Economics, Vol. 87, No. 3, pp. 355–374.

Deepfake, Deep Scars-Protecting Children's Rights in the Age of Artificial Intelligence

Sophie Kung, Nicole Lin, Joy Hong, Peter Lee, Yaping Tsao
Taipei Municipal Zhong Shan Girls High School, Taipei, Taiwan

Abstract

This paper examines the escalating threat of AI-generated deepfake abuse against minors, focusing on the systemic dynamics that sustain and amplify these harms in the contemporary digital landscape. Adopting a systems-thinking approach, the research utilizes analytical tools such as the Iceberg Model and Causal Loop Diagrams to move beyond an event-oriented perspective of individual tragedies—such as the high-profile Elijah Heacock case—to a structural analysis of the digital environment. The methodology incorporates a comparative legislative review of frameworks in Taiwan, the United States, the European Union, and South Korea, alongside quantitative data synthesis from reporting trends provided by the National Center for Missing & Exploited Children (NCMEC) and the Internet Watch Foundation (IWF). The core analysis identifies four primary reinforcing loops that drive the deepfake crisis: the Low-Barrier Reinforcing Loop, the Anonymity–Impunity Reinforcing Loop, the Normalization–Rationalization Reinforcing Loop, and the Personalized Exploitation Reinforcing Loop. Empirical evidence indicates that AI-generated sexual exploitation is an urgent global concern, with estimates suggesting over 1.2 million children were affected by "nudified" deepfake tools in 2025 alone. The paper concludes by presenting a "Solution Landscape" that integrates technological, legal, and ethical interventions, including "Digital DNA" for content authentication and "Mandatory Empathy Windows" to introduce psychological friction in generative processes. By addressing the deep-seated structures that normalize digital harm, this research advocates for a coordinated global response to safeguard children's rights in the age of artificial intelligence.

Keywords: Deepfake technology, Child protection, Artificial intelligence, Systems thinking, Digital rights.

Paper type: Academic Research Paper.

1 Introduction

On February 28, 2025, a 16-year-old boy named Elijah Heacock took his own life after becoming the target of an AI-based sextortion scheme. Elijah had a loving family, supportive friends, and a bright future; yet a fabricated nude image of him, generated through artificial intelligence, was used to threaten, humiliate, and blackmail him. The perpetrator demanded money in exchange for not sending the image to his family and friends. Although the image was fake, the fear, shame, and emotional devastation it caused were entirely real. Elijah's death is not only a personal tragedy but also a warning of how emerging technologies can be weaponized against children and adolescents in deeply harmful ways.

Artificial intelligence has transformed modern society at an extraordinary pace, creating new possibilities across education, business, and daily life. However, AI is also a double-edged sword. Among the most alarming examples is deepfake technology, which enables the creation of highly realistic but fabricated content with minimal cost. When such technology is used against minors, the consequences are severe, violating their privacy, dignity, and basic rights. To examine this issue beyond individual incidents, this report adopts a systems-oriented approach combining structural analysis, comparative legal review, and case-based evidence.

2 Methodology and Positionality

2.1 Systems Thinking Approach

This report adopts a systems-thinking approach to examine how AI-generated deepfake abuse is produced, sustained, and intensified across digital environments. It uses analytical tools such as the Iceberg Model and Causal Loop Diagrams to identify the structural conditions, feedback processes, and psychological assumptions that underlie visible incidents. By shifting the analysis from an event-oriented perspective to a systemic one, the report traces how technological accessibility, regulatory delay, and offender motivations interact to perpetuate harm.

The research draws on three main approaches:

- **Comparative Legislative Analysis:** Examining legal and policy frameworks in Taiwan, the United States, the European Union, and South Korea to identify governance gaps.

- **Quantitative Data Synthesis:** Analyzing reporting trends and child-protection statistics from NCMEC and the IWF to assess escalation patterns.
- **Case Study Review:** Connecting system-level analysis with lived human consequences through incidents like the Elijah Heacock case and the Westfield High School case.

2.2 Positionality

The authors are three high school students from Taiwan, a society with high internet penetration. As digital natives, they occupy a dual position: they are both researchers and members of the population most exposed to AI-enabled harm. This positionality provides a unique perspective on peer-related digital risks, recognizing that these issues are intimately embedded in the everyday digital environments of young people. To ensure objectivity, the analysis is grounded in international data and cross-national cases.

3 Problem Statement

The misuse of AI-generated deepfake content has become a serious threat to children's rights. Deepfake technology first emerged in 2017 and rapidly evolved from fringe face-swapping into a broader social threat. By the 2020s, it had become increasingly linked to child sexual exploitation and financially motivated sextortion.

3.1 Historical Context

Deepfake technology transitioned from isolated incidents to systemic risks. In 2021, the Westfield High School case in New Jersey saw students use AI to generate explicit images of female classmates, highlighting the peer-to-peer nature of the threat. Between October 2021 and March 2023, the FBI received over 13,000 reports of online financial sextortion involving minors, linked to at least 20 suicides. The IWF reported action on 312,030 reports of child sexual abuse material (CSAM) in 2025, a 7% increase from 2024.

3.2 Impact and Significance

Recent European Parliament analysis indicates that pornographic material accounts for approximately 98% of deepfakes, with global sharing projected to rise to 8 million by 2025. UNICEF reports that in a survey covering 11 countries, at least 1.2 million children were estimated to have had their images "nudified" through deepfake tools in the previous year. For minors, these harms result in severe emotional consequences, including anxiety, depression, and suicidal ideation. The "permanence" of digital data means that even if a deepfake is proven false, the digital shadow can follow a victim indefinitely.

3.3 Projection and Escalation

To illustrate the possible scale of future risk, we present a scenario projection based on NCMEC reports (4,700 in 2023; 67,000 in 2024). The model applies a hypothetical annual growth rate of 150%, assumption-based on the explosive rise observed in recent years.

The projection uses the compound-growth formula: $\mathbf{N(t) = N_o(1 + r)^t}$.

Table 1: Scenario Projection of AI-Related Abuse Reports

Year	Projected Reports
2025	167,500
2026	418,750
2027	1,046,875
2028	2,617,188

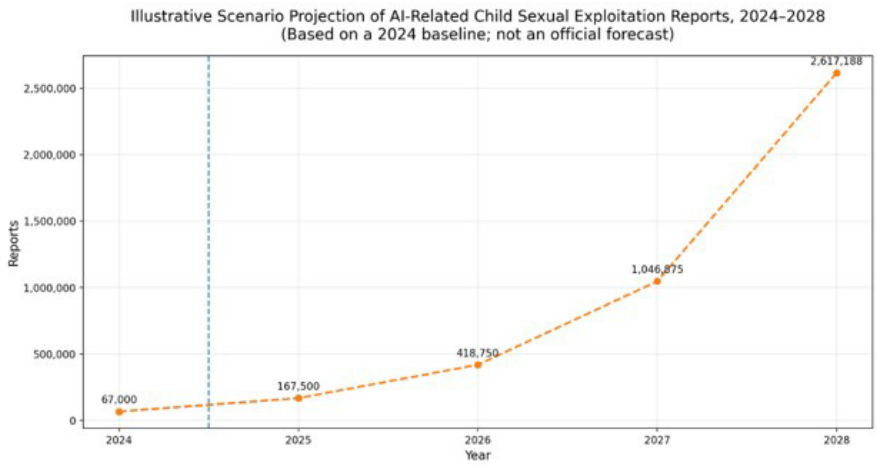


Figure 1. Illustrative scenario projection of AI-related child sexual exploitation reports, 2024–2028.

4 Systems Mapping

4.1 The Iceberg Model

To understand why this problem persists, we distinguish between visible events and underlying drivers:

- **Events:** Individual suicides, sextortion cases, and viral fake images.
- **Patterns:** Recurring trends of cross-jurisdictional offending and the rapid speed of digital circulation.
- **Structures:** Technological accessibility, fragmented legal systems, and platform anonymity.
- **Mental Models:** Harmful beliefs that "synthetic images cause no real harm" or that digital anonymity provides absolute legal protection.

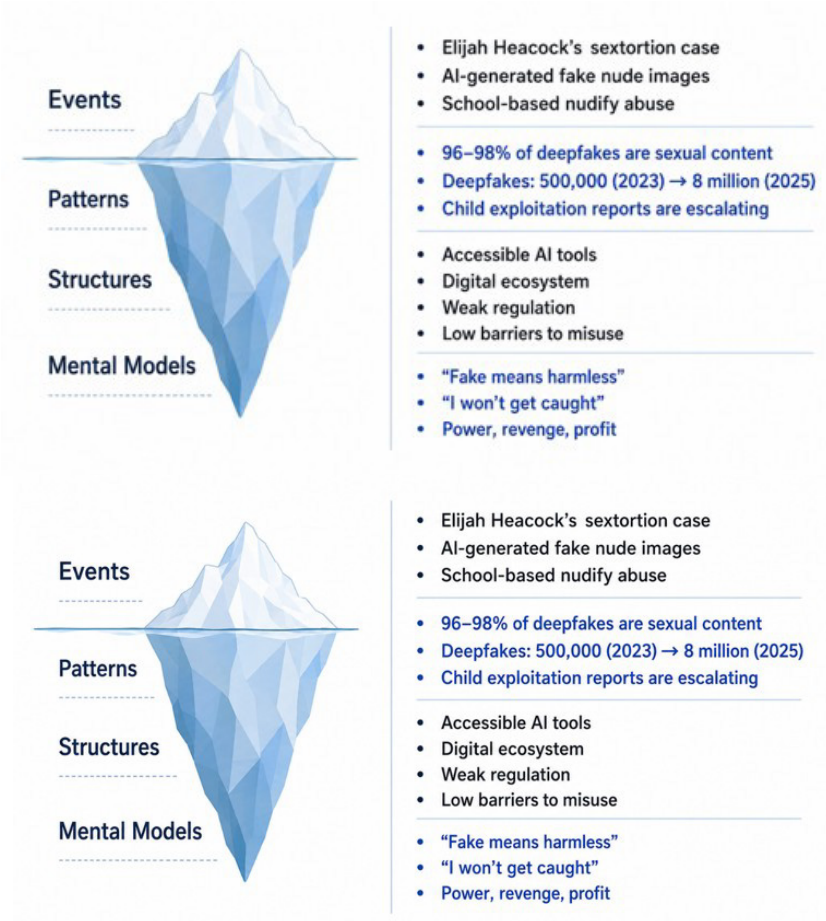


Figure 2. Iceberg model of AI-generated deepfake abuse against minors.

4.2 Causal Loop Reinforcing Logic

The deepfake crisis is driven by self-reinforcing dynamics:

- **Anonymity–Impunity Loop:** Concealment mechanisms allow offenders to obscure identities across jurisdictions. This reduces traceability, lowering the likelihood of sanctions and decreasing the psychological cost of offending.
- **Normalization–Rationalization Loop:** As more abuse occurs, desensitization follows. Increased exposure weakens moral resistance, treating abuse as routine or socially tolerated.
- **Personalized Exploitation Loop:** Reusable knowledge and access to victim-specific data increase the effectiveness of targeted abuse over time.

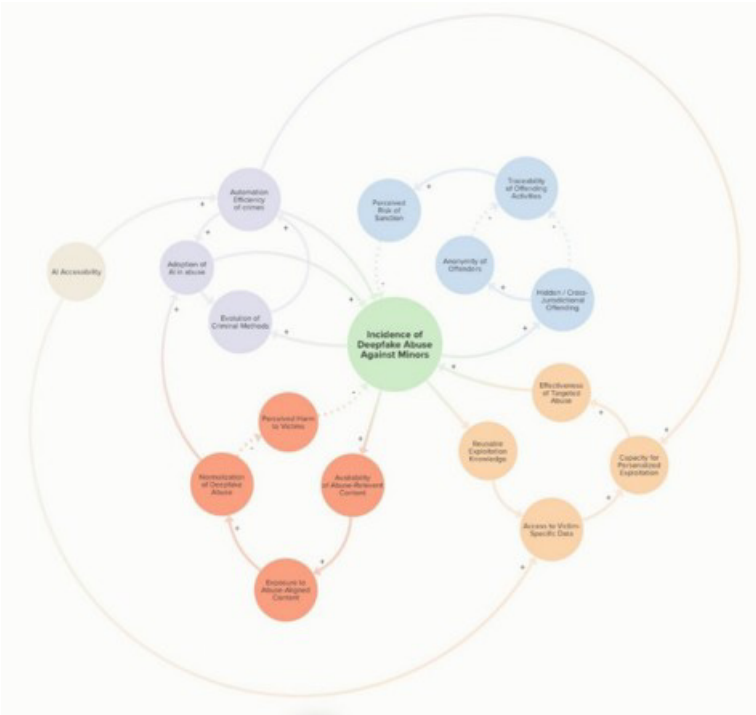


Figure 3. Causal loop diagram of the systemic dynamics sustaining AI-generated deepfake abuse against minors.

5 Comparative Legislative Analysis

5.1 Taiwan

Taiwan's **Child and Youth Sexual Exploitation Prevention Act (2024)** explicitly includes sexual imagery of minors produced through computer synthesis or other technological methods. While it recognizes synthetic content as exploitation, enforcement against overseas servers remains a primary challenge.

5.2 United States

The U.S. has introduced the **DEFIANCE Act** and the **Take It Down Act (2024)**, aiming to provide victims with civil causes of action and mandated removal from platforms. However, protection remains decentralized and varies by state.

5.3 European Union

The **EU AI Act** categorizes deepfake generation as "limited risk," requiring clear disclosure/labeling. While GDPR provides strong privacy protections, criminal law harmonization for synthetic content is still evolving.

5.4 South Korea

Following a 2024 scandal, South Korea passed a bill to criminalize the **possession or viewing** of sexually explicit deepfakes. This represents a shift toward punishing consumption to dismantle the market for synthetic abuse.

6 Solution Landscape

6.1 Solution 1: Mandatory Empathy Window

This intervention introduces psychological friction into the generative process. AI models would be required to pause and display the potential real-world harm of a prompt before generating realistic human imagery, triggering moral hesitation in the offender.

6.2 Solution 2: Digital DNA (Provenance)

Targets the Anonymity Loop by increasing traceability. Utilizing the **C2PA standard**, every generated file would contain persistent evidence of its origin, recording the tool, timestamp, and potentially user credentials.

6.3 Solution 3: Digital Bloodhound (AI Monitoring)

Leverages AI as a tool for detection rather than concealment. Specialized AI "bloodhounds" would be deployed to detect subtle pixel or movement anomalies that reveal manipulated content across platforms.

6.4 Solution 4: Jamming Signals (Image Cloaking)

Empowers individuals via tools like "Glaze" or "Nightshade," which apply invisible perturbations to images. These perturbations make raw data unusable for AI models trying to learn a person's likeness, sabotaging criminal efficiency at the source.

Table 2. Summary of proposed system-level interventions for AI-generated deepfake abuse against minors.

Solution	Core idea	Main target	Expected effect
Mandatory Empathy Window	Interrupt harmful sharing by forcing offenders to confront real human suffering before they share deepfake sexual content.	**Normalization–Rationalization Reinforcing Loop**	Reduces numbness, disrupts group-think, and weakens the belief that synthetic abuse is harmless.
Digital DNA	Embed traceable origin information into AI-generated files so content can be linked back to its source.	**Anonymity–Impunity Reinforcing Loop**	Restores accountability, increases traceability, and raises the perceived risk of sanction.
Digital Bloodhound	Use AI detection tools and real-time cross-border cooperation to identify and block harmful synthetic content.	**Anonymity–Impunity Reinforcing Loop**	Makes harmful content easier to detect and reduces offenders' ability to hide across platforms and jurisdictions.
Jamming signals	Add invisible perturbations or anti-scraping signals to personal images so AI systems cannot easily interpret or reuse them.	**Personalized Exploitation Reinforcing Loop**	Reduces the machine-readability of victim data and weakens the efficiency of personalized exploitation.

7 Feasibility and Risk Assessment

7.1 Feasibility

The C2PA standard is already supported by industry leaders like Google and Microsoft. However, a significant barrier is the **"Open-Source Gap"**. Offline, "jailbroken" models can bypass built-in empathy windows or watermarking requirements.

7.2 Privacy vs. Protection

There is an inherent tension between traceability (Digital DNA) and digital anonymity. Any solution must include strong legal guardrails to ensure these tools are not weaponized by authoritarian regimes against dissidents.

8 Conclusions

The rise of AI-generated content presents a dual crisis: unprecedented creative possibility versus systemic risks to the fundamental rights of children. Deepfake abuse is not merely a collection of isolated events, but a systemic problem driven by technological ease and legal fragmentation. Protecting the next generation requires a shift from "reactive removal" to "proactive prevention" through an integrated approach of technology (Digital DNA), law (international coordination), and social behavior (Empathy Windows).

9 Ethics declaration

This research was conducted according to ethical guidelines for child protection. All case studies (e.g., Elijah Heacock) utilize publicly available reports, and the authors have prioritized the dignity of victims throughout the process.

10 AI declaration

Generative AI tools were utilized to synthesize reporting trends and to generate conceptual diagrams (the Iceberg Model and Causal Loop Diagrams). These tools were used strictly for structural visualization and data synthesis to enhance the clarity of the systems-thinking methodology.

References

Bellas, G. (2025, March). Deepfakes in the courtroom: Problems and solutions. Illinois State Bar Association. https://www.isba.org/sections/bench/newsletter/2025/03/deepfakesinthecourtroomproblemsandsolutions

Cable News Network. (2024, September 26). South Korea to criminalize watching or possessing sexually explicit deepfakes. https://edition.cnn.com/2024/09/26/asia/south-korea-deepfake-bill-passed-intl-hnk

C2PA (Coalition for Content Provenance and Authenticity). (2024). Content credentials: Strengthening multimedia integrity in the generative AI era. https://c2pa.org/wp-content/uploads/sites/33/2025/10/content_credentials_wp_0925.pdf

Canadian Broadcasting Corporation. (2023). Calgary teen accused of using AI to sexualize photos of high school girls. https://www.cbc.ca/news/canada/calgary/ai-sexualized-photos-teen-charged-9.7001828

CBS News. (2024). A teen died after being blackmailed with AI-generated nudes. His family is fighting for change. https://www.cbsnews.com/news/sextortion-generative-ai-scam-elijah-heacock-take-it-down-act/

Chandra, B., Dunietz, J., Roberts, K., Lee, Y., Fontana, P., & Awad, G. (2024). Reducing risks posed by synthetic content: An overview of technical approaches to digital content transparency (NIST AI 100-4). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.AI.100-4

Coalition for Content Provenance and Authenticity. (n.d.). C2PA: Verifying media content sources. https://c2pa.org/

College of Policing. (n.d.). Evolution of AI in child sexual abuse material. https://www.college.police.uk/article/evolution-ai-child-sexual-abuse-material

Eithos. (n.d.). Deepfake technology: History useful to know. https://eithos.eu/deepfake-technology-1-history-useful-to-know/

Eucrim. (2025). Hit against AI-generated child sexual abuse material (Operation Cumberland). https://eucrim.eu/news/hit-against-ai-generated-child-sexual-abuse-material/

Europol. (2024). Internet organised crime threat assessment (IOCTA) 2024. https://eucrim.eu/news/iocta-2024/

Europol. (2025, February 28). 25 arrested in global hit against AI-generated child sexual abuse material. https://www.europol.europa.eu/media-press/newsroom/news/25-arrested-in-global-hit-against-ai-generated-child-sexual-abuse-material

Europol. (2025). Operation Cumberland: Multilateral action against online child exploitation. https://www.europol.europa.eu/newsroom

European Parliament. (2025). Children and deepfake: Risks, impacts and policy responses. https://www.europarl.europa.eu/RegData/etudes/BRIE/2025/775855/EPRS_BRI(2025)775855_EN.pdf

European Union. (2024). Artificial Intelligence Act (Article 50). https://artificialintelligenceact.eu

Federal Bureau of Investigation. (n.d.). Sextortion: A growing threat targeting minors. https://www.fbi.gov/contact-us/field-offices/nashville/news/sextortion-a-growing-threat-targeting-minors

Francisco, S. M., Ferreira, P. da C., Veiga Simão, A. M., & Salgado Pereira, N. (2023). Measuring empathy online and moral disengagement in cyberbullying. Frontiers in Psychology, 14, Article 1061482. https://doi.org/10.3389/fpsyg.2023.1061482

Frontiers in Psychology. (2026). Algorithmic seduction: Ethical boundaries in AI-powered consumer nudging. https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2026.1675412/full

Incident Database. (2025). Voices of the Strong 44 incident. https://incidentdatabase.ai/cite/1350/

Internet Watch Foundation. (2024). AI becoming a “child sexual abuse machine” adding to record levels of online abuse. https://www.iwf.org.uk/news-media/news/ai-becoming-child-sexual-abuse-machine-adding-to-dangerous-record-levels-of-online-abuse-iwf-warns/

Internet Watch Foundation. (2024). Annual data and insights report 2024. https://www.iwf.org.uk/annual-data-insights-report-2024/data-and-insights/reports-assessment/

Internet Watch Foundation. (2024). Glossary: Terms and acronyms. https://www.iwf.org.uk/annual-data-insights-report-2024/glossary/

INTERPOL. (2024). Beyond illusions: Exploring synthetic media and its implications for law enforcement. https://www.interpol.int/en/content/download/21179/file/BEYOND%20ILLUSIONS_Report_2024.pdf

Jahn, L., Rendsvig, R. K., Flammini, A., Menczer, F., & Hendricks, V. F. (2025). A perspective on friction interventions to curb the spread of misinformation. npj Complexity, 2, Article 31. https://doi.org/10.1038/s44260-025-00051-1

KSAT. (2023, December 2). Teen girls are being victimized by deepfake nudes, one family is pushing for more protections. https://www.ksat.com/gallery/news/2023/12/02/teen-girls-are-being-victimized-by-deepfake-nudes-one-family-is-pushing-for-more-protections/

Lennon, C. (2026, January). AI enables new forms of child exploitation, UN warns. United Nations. https://news.un.org/en/story/2026/01/1166827

National Center for Missing & Exploited Children. (2024). CyberTipline report data. https://www.missingkids.org/gethelpnow/cybertipline/cybertiplinedata

National Center for Missing & Exploited Children. (2024). Generative AI CSAM is CSAM.

National Center for Missing & Exploited Children. (2024). Insights from NCMEC’s 2023 CyberTipline report. https://www.missingkids.org

National Center for Missing & Exploited Children. (2025). 2024 in numbers. https://www.missingkids.org

National Center for Missing & Exploited Children. (2025). The deepfake dilemma: New challenges protecting students’ confidentiality. https://www.missingkids.org/blog/2025/the-deepfake-dilemma-new-challenges-protecting-students-confidentiality

National Center for Missing & Exploited Children. (2026). Generative AI (GAI): Issues brief. https://www.missingkids.org

National Center for Missing & Exploited Children. (2026). The work never stops: A first look at NCMEC’s 2025 data.

Negreiro Achiaga, M. del M. (2025). Children and deepfakes. European Parliamentary Research Service, European Parliament.

NIST (National Institute of Standards and Technology). (2024). Response to AI transparency and provenance standards. https://downloads.regulations.gov/NIST-2024-0001-0030/attachment_1.pdf

Sample, I. (2018, November 12). What are deepfakes – and how can you spot them? The Guardian. https://www.theguardian.com/technology/2018/nov/12/deep-fakes-fake-news-truth

Shan, S., Cryan, J., Wenger, E., Zheng, H., Hanocka, R., & Zhao, B. Y. (2023). Glaze: Protecting artists from style mimicry by text-to-image models. USENIX Security Symposium. https://www.usenix.org/conference/usenixsecurity23/presentation/shan

Shan, S., Wenger, E., Zhang, J., Li, H., Zheng, H., & Zhao, B. Y. (2020). Fawkes: Protecting personal privacy against unauthorized deep learning models. In 29th USENIX Security Symposium (USENIX Security 20) (pp. 1589–1604). USENIX Association. https://www.usenix.org/system/files/sec20-shan.pdf

UK Ministry of Justice. (2025). Crime and policing bill 2025: Child sexual abuse material factsheet. https://www.gov.uk/government/publications/crime-and-policing-bill-2025-factsheets/crime-and-policing-bill-child-sexual-abuse-material-factsheet

United Nations Children’s Fund. (n.d.). Deepfake abuse is abuse. https://www.unicef.org/press-releases/deepfake-abuse-is-abuse

United Nations Children’s Fund. (2024). Artificial intelligence and child sexual abuse and exploitation. https://www.unicef.org/media/178571/file/UNICEF%20AI%20CSEA%20Brief_2.pdf

United Nations Office on Drugs and Crime. (n.d.). Online child sexual exploitation and abuse. https://www.unodc.org/cld/zh/education/tertiary/cybercrime/module-12/key-issues/online-child-sexual-exploitation-and-abuse.html

UNICEF. (2026). Artificial intelligence and child sexual abuse and exploitation (Issue brief). https://www.unicef.org/reports/artificial-intelligence-and-child-sexual-abuse-and-exploitation

U.S. Department of Homeland Security. (2024). Generative AI bulletin. https://www.dhs.gov/sites/default/files/2024-09/24_0920_k2p_genai-bulletin.pdf

U.S. Department of Justice. (2023). National strategy for child exploitation prevention and interdiction. https://www.justice.gov/criminal-ceos/2023-national-strategy-child-exploitation-prevention-interdiction

VentureBeat. (2024). Glaze 2: New version of anti-AI scraping tool for artists launches. https://venturebeat.com/ai/glaze-2-new-version-of-anti-ai-scraping-tool-for-artists-launches/

全國法規資料庫。 (n.d.)。 人工智慧基本法。 https://law.moj.gov.tw/LawClass/LawAll.aspx?pcode=H0160093

國家發展委員會，& 理律法律事務所。 (n.d.)。 人工智慧之相關法規國際發展趨勢與因應策略研究計畫報告。 https://ws.ndc.gov.tw/Download.ashx?u=LzAwMS9hZG1pbmlzdHJhdG9yLzEwL3JlbGZpbGUvNTc0NC8zMTk5OC8xY2MwZWFiZi1jNTMzLTQ0M2QtYjc1Mi1iM2U4MDkwZmYzMTIucGRm&n=44CM5Lq65bel5pm65oWn5LmL55u46Zec5rOV6KaP5ZyL6Zqb55m85bGV6Lao5Yui6IiH5Zug5oeJ44CN5aeU6KiX56CU56m26KiI55Wr57WQ5qGI5aCx5ZGKLnBkZg%3d%3d&icon=.pdf

Sexual Violence Towards Female Domestic Migrant Workers

Kai-En Chang, Yo-Yo Chen, Yun-Hsuan Chu
Taipei Municipal Song-Shan Senior High School, Taipei, Taiwan

Abstract

In Taiwan, one of six female migrant workers has been sexually assaulted. And this phenomenon not only occurs in Taiwan but the whole world. UN WOMEN defines that nearly half of the world's 304 million international migrant workers are women (migration data portal, May 14 ,2025), crucially 765 of the 75.6 million domestic workers worldwide are female (UN women, 2025, 23 October). Because these women are primarily employed as domestic workers or caregivers, they operate within private households. This kind of private areas mean they gain less visibility and protection, exposing them to a higher risk of suffering from exploitation, abuse, or even sexual violence compared to non-female migrant domestic workers. A research by Linn et al. (2024) indicates that approximately 60 per cent of workers surveyed had experienced violence within the previous 12 months, which is markedly higher than the percentages of women who had reportedly experienced violence in their lifetime in Myanmar (20.6 per cent) and in Thailand (15.4 per cent). Our research analyses different aspects, including legislation, racial hierarchy and so on. After evaluating existing solutions' pros and cons, we came up with some potential solutions in the hope that this problem can be recognized and ameliorated.

Keywords: Systemic Power Imbalance, Human Rights Fallacy, Female domestic migrant worker, Global and Local Racial Hierarchies, Intermediary System

Paper type: Practical Paper

1 Challenges Landscape

The suffering of sexual violence stems from two dimensions: root causes, which weaken the position of female migrant workers, and perpetuating factors, which hinder resolution. Legally, linguistically, and economically disadvantaged, these workers often face societal discrimination and indifference (CEDAW, 2008). This marginalization discourages victims from seeking help and complicates the delivery of assistance.

Furthermore, the concentration of female migrant workers in private caregiving environments increases their vulnerability to harassment and rape (The Control Yuan, 2022). Seeking justice is further obstructed by the risk of job loss or deportation, alongside significant loopholes in law enforcement (AIT, 2024; 2025). This exploitation chain is often completed by intermediaries who force reconciliations and conceal abuse cases, ensuring the difficulties faced by these women remain unrecognized and unresolved (The Control Yuan, 2025).

1.1 Systemic Power Imbalance

Systemic power imbalance means one group consistently holds more economic, legal, and social power over another. Sexual violence against female migrant workers is not merely the result of individual misconduct but is rooted in systemic power imbalances shaped by patriarchal norms and reinforced by high-risk labor and

migration structures (CEDAW, 2008). As women, also as migrant workers, they are often vulnerable to exploitation (The Control Yuan, 2025). Women are often objectified, not considered as a human but as sexual value (Taiwan Taipei District Court, 2024). And as migrant workers, they are usually under the pressure of debt or the demand for a job, which puts them in a weaker position (American Institute in Taiwan, 2024). Most female migrant workers are employed as care giver, working in private homes means they receive less visibility and protection, which is likely to exposes them more to exploitation and even sexual violence by employers or other people in the household (ILO, 2011, December 15). Reporting is difficult for them too because they might lose their job, which is a risk few would take. As migrant workers, they will also have to face the barrier of language, which, combined with the dilemma they are already facing, is likely for them to bear and endure (Ministry of Health and Welfare, 2024).

1.2 Human Rights Fallacy

1.2.1 Neglect of Negative Human Rights for Migrant Workers

A prevailing bias reduces migrant workers to "labor commodities," misconstruing fundamental human rights as "extra benefits" or "privileges". When employers view workers as "private property" rather than legal

subjects, basic rights become a matter of "benevolence" rather than statutory protection. However, human dignity should never be weighed against economic profit (Rothna Begum / Human Rights Watch,2015)

The tension between Perceived Resource Costs and Legislative Prioritization can be explained through "Bounded Rationality" (Simon, 1955) and "Mental Accounting" (Thaler, 1999). Due to cognitive limits, the public often conflates essential protections with "expensive" extras, leading legislators to de-prioritize these issues in favor of more "urgent" matters. Ultimately, as Kapp (Kapp, 1950) argued, neglecting these marginalized needs forces society to pay a much higher social cost later to address crime and welfare systemic failures.

1.3 Global and Local Racial Hierarchies

The third root cause is the pervasive racial hierarchy and dehumanization inherent in global migration. Developed nations often relegate labour from developing countries to "second class" status. As Sassen (Sassen ,1991) observed, "Global Cities" depend on marginalized migrant groups while systematically excluding them from legal and social protections. In Taiwan, this manifests as a collective bias viewing citizenship as inherently superior to Southeast Asian migrants.

Such cultural bias facilitates sexual violence through Dehumanization; when workers are viewed as "disposable tools" rather than individuals, perpetrator empathy diminishes (Anderson, 2000). Furthermore, Taiwanese social discourse often "accepts" this violence as it aligns with a racialized logic placing Southeast Asian bodies at the social margins (Lan, 2006). This psychology catalyses the Social Apathy Loop, where the suffering of female workers is dismissed as a systemic "norm".

1.3.1 The Cycle of Social Apathy and Complicit Silence

Dehumanization fosters a Complicated Silence within social support systems. Due to hierarchical prejudice, the public often misclassifies abuse as a "private labour dispute" rather than a human rights violation. As Parreñas (Parreñas ,2015) notes, a "hierarchy of citizenship" creates a power vacuum at the bottom of the global care chain, where migrant workers are systematically denied full legal protection.

This apathy directly reinforces the Silence and Endurance of victims. According to a 2024 SPA survey, 1 in 6 female migrant workers in Taiwan has experienced gender-based violence, yet most remain silent (SPA, 2024). Perpetrators exploit this vulnerability by using threats of "immediate deportation" and operating under systemic impunity, emboldened by the knowledge that social biases will likely discredit the victim's voice.

1.3.2 The Visibility Trap: Low Reporting Rates and Statistical Fog

The sum of these factors results in a chronic Low Reporting Rate, creating a dangerous Visibility Trap. Official statistics often fail to reflect the true scale of the crisis because the costs of reporting—including job loss, debt bondage from high brokerage fees, and facing a judicial system lacking cultural sensitivity—far outweigh the hope of seeking justice (U.S. Department of State, 2024).

When the "visibility of actual occurrence" in official records remains low, it leads to a decline in social awareness. News media and the public encounter such information less frequently, leading to further estrangement and neglect of the issue. Without public oversight, the pressure on the government to change policy diminishes, which in turn justifies the continued existence of "high-risk labor structures." This lack of momentum for reform reinforces the original sense of racial superiority; the absence of public protest further cements the Subaltern

status of migrant workers, allowing the vicious cycle of sexual violence to continue indefinitely.

1.4 Environment of Impunity

Beyond the root causes, several perpetuating factors intertwine to form a "vicious closed loop”, ensuring that sexual violence against female migrant workers is not only possible but systematically sustained over the long term

1.4.1 Systemic Impunity and the Environment of Low Perceived Risk

The primary factor in the persistence of violence is the "low-risk environment" perceived by perpetrators. This sense of impunity stems from the failure of administrative oversight and the inefficiency of legal enforcement. An extreme example is the 2024 "Keelung Case," in which a wanted sex offender successfully passed government screening to legally hire a foreign caregiver, subsequently subjecting her to nine months of abuse—during which the local government failed to conduct a single site visit(Zeng, 2025).When perpetrators observe loose social monitoring and massive loopholes in legal execution, violent behavior persists because it is

seen as "cost-free." This environment reinforces illicit behavior and further erodes the deterrent power of the law.

1.4.2 Privacy as a Shield for Violence: The Invisibility of Domestic Space

The nature of the workspace profoundly affects the frequency of violence. As Parreñas (Parreñas, 2015) observes, employment within private domestic spaces severs social oversight. In Taiwan, this physical isolation causes violence to shift into hidden environments where public apathy treats domestic occurrences as "private matters". This concealment not only facilitates abuse but also obstructs evidence collection, fracturing judicial remedies.

Furthermore, this isolation creates a "Data & Visibility Gap." When violence remains confined to private spheres and victims remain silent, the resulting lack of official statistics misleads policymakers into believing the issue is managed. This statistical invisibility subsequently perpetuates systemic withdrawal, stratification, and continued economic oppression.

1.4.3 Hierarchy of Citizenship and the Power Vacuum

Parreñas further notes the existence of a global "hierarchy of citizenship," where the laws of developed nations often view migrant workers as "non-citizens" or merely "temporary labor"(Parreñas, 2015).In Taiwan, this hierarchy translates into a power vacuum—a dangerous space where "the law cannot reach, outsiders cannot see, and the perpetrator has the final say." Legal systems are significantly biased toward employers. Because migrant workers lack full citizenship rights and the freedom to change employers, they are trapped in a high-risk structure built on prejudice. This structural inequality grants employers' absolute power, transforming the labor-management relationship into one oppression.

1.4.4 Economic Disempowerment and the Cost of Refusal

The systemic exploitation of migrant workers significantly increases their "cost of refusal." High brokerage fees create a state of "debt bondage" from the onset of their labor in Taiwan (Workforce Development Agency, 2021). Simultaneously, these women often bear the burden of "transnational motherhood," serving as the primary economic pillars for their families back home (Parreñas, 2015).

Furthermore, Southeast Asian gender norms regarding chastity impose a "family disgrace" on victims of sexual violence, regardless of consent. This cultural stigma compels victims to choose silence to avoid familial disdain. Caught between "personal safety" and "family livelihood," workers often prioritize debt repayment over seeking justice. This induced "silence and endurance" results in low reporting rates, ultimately reinforcing the invisibility of the entire system.

1.5 Legislative Loopholes and Administrative Negligence

1.5.1 Failure of Administrative Due Process

The Control Yuan's reports highlight a critical failure in inter-agency cooperation; administrative bureaucracy often stalls case transfers, leading to "dark figures" of unreported abuse. A notable instance involved a case suppressed for two years, only surfacing via a whistleblower petition. By the investigation's conclusion, the victim had already returned home without compensation, bearing the psychological and financial scars alone.

Evidence of falsified records further underscores this systemic negligence. Municipal governments frequently fail to conduct mandatory on-site inspections, leaving discrepancies between actual and official residences unnoticed. Furthermore, bypassed departure verification procedures allow employers to forcibly deport workers before grievances are settled. These failures render the "Manpower Agency Evaluation System" futile; agencies concealing abuse have maintained "Grade B" ratings, proving the system is fundamentally disconnected from reality.

1.5.2 Legal Loopholes and Deficient Protection Mechanisms

The current legal framework lacks clear penalties for local governments regarding administrative negligence, such as mishandled case transfers or falsified records. This legislative vacuum leaves migrant workers without protection when their fundamental rights are violated. A primary concern is the "window period" during an employer transfer; although a de facto employment relationship exists, the lack of formal registration allows employers to evade statutory obligations, significantly increasing the worker's vulnerability. Furthermore, the 1955 hotline referral mechanism typically assigns cases to the authority where the employer is registered rather

than where the incident occurred. This jurisdictional ambiguity leads to bureaucratic "buck-passing," causing urgent cases to be indefinitely shelved.

1.6 Pressure from Intermediary Companies and Structural Exploitation

1.6.1 Forced Settlements: Evidence from the Keelung and Chiayi Cases

In the global labor migration system, manpower agencies act as the primary gatekeepers, controlling everything from daily life and work assignments to communication with employers. For newly arrived workers, the agency is often their sole point of contact. Due to language barriers and a lack of information, many remain unaware of alternative support channels. Even for experienced workers, agencies still wield immense power over their legal status and livelihood.

The actions of an agency can determine whether a case of sexual assault or violence is resolved or suppressed. In the Chiayi case, the agency failed to report the incident and illegally coerced the victim into a settlement. The Ombudsman's report reveals that the agency pressured the victim to accept NT$100,000 and sign a settlement agreement. By attempting to "bypass" legal procedures to minimize their own financial risks and administrative burden, the agency effectively silenced the victim and became a primary catalyst for the delay of justice. (The control Yuan, 2025, August 22)

1.6.2 International Context: The Kafala System and Total Control

In the Gulf Cooperation Council (GCC) countries, the Kafala (Sponsorship) System represents a more extreme form of labor control than Taiwan's framework. Under Kafala, employers exercise near-total authority by withholding passports and controlling visas, effectively stripping workers of their freedom of movement. Bajracharya and Sijapati (Bajracharya, R., & Sijapati, B. ,2012) report that authorities often forcibly return fleeing workers, while employers may retaliate against those escaping abuse by filing false "absconding" charges.

While Taiwan has made progress, these global comparisons underscore that systemic violations remain a persistent threat. Female migrant workers worldwide endure multifaceted exploitation—from wage discrimination to physical violence. For many, resistance remains silent and invisible, making their protection a critical human rights priority for the international community.

2 Systems Thinking Approach

2.1 Systems Map

Below is a systems map that we consolidated our challenge landscape into; we organized four loops (Figure 1)The blue one is the social cultural loop, the orange one is the institutional loop, the green one is the report procedure loop, the yellow one is the mental model loop, and the purple one is the structural dependency loop.

Figure 1: System Map

2.2 Systemic Loop Analysis

2.2.1 Socio-Cultural & Institutional Loops

The Social-Cultural Loop illustrates how racial discrimination evolves into the dehumanization and objectification of female migrant workers, fostering a "structural complicity of apathy". This environment compels victims to endure abuse in silence, leading to low reporting rates that further entrench racial biases.

Simultaneously, the Institutional Loop highlights how legislative loopholes and systemic impunity lower the perceived risk for perpetrators. The resulting physical isolation within private homes creates a "visibility gap," obstructing law enforcement and stabilizing the environment of impunity.

2.2.2 Reporting & Mental Model Loops

The Report Procedure Loop reveals a failing protection system where hotlines fail to address jurisdictional conflicts or workplace discrepancies. These failures lead to administrative delays and falsified inspection records, creating a "legal vacuum" where formal protections exist but are inaccessible. Furthermore, the Mental Model Loop identifies how intermediaries prioritize profit and employer relationships over worker safety. By forcing coerced settlements, intermediaries create severe information asymmetry, distorting evaluation systems with falsified data.

2.2.3 Structural Dependency Loop

The Structural Dependency Loop highlights how legal and financial vulnerabilities—exacerbated by debt bondage—strip workers of bargaining power. This dependency, intensified by spatial isolation and information asymmetry, makes it nearly impossible for workers to leave abusive employers, thereby increasing the probability of sustained violence. Collectively, these loops form a self-reinforcing system that perpetuates sexual violence against female migrant workers.

2.2 Iceberg Model

Below is our iceberg model (Figure 2)

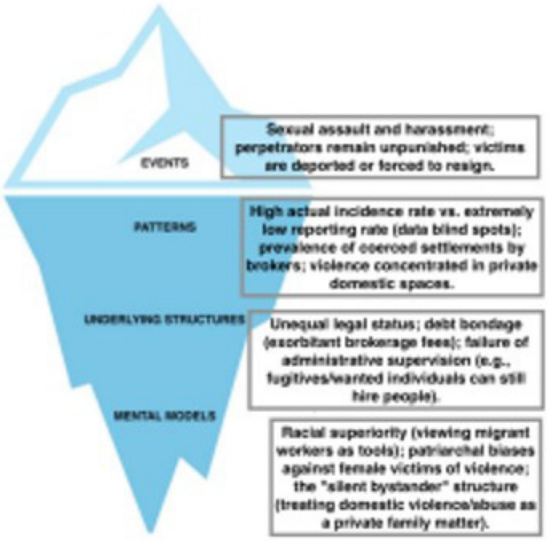


Figure 2: Iceberg Model

2.3 Stakeholders' Map

This map shows stakeholders who either exacerbate or ameliorate the problem of sexual violence towards female migrant workers (Figure 3) .

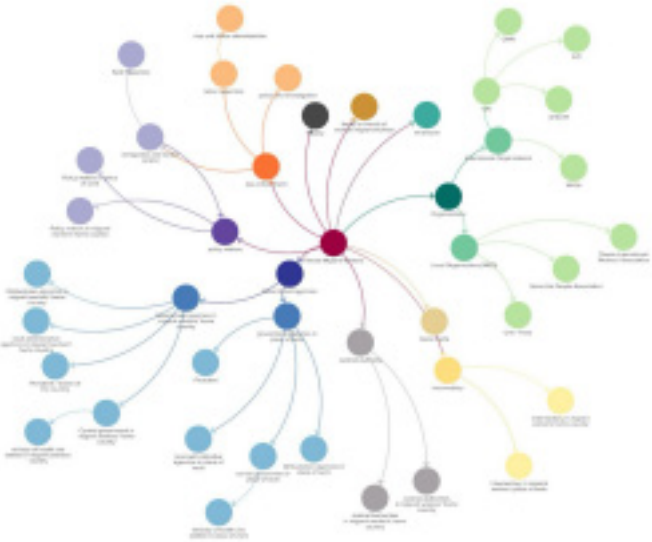


Figure 3: Stakeholders' Map

3 Solutions Landscape

3.1 Disadvantages of Existing Solutions

3.1.1 International Frameworks

- UN Resolution 78/180 (2023): While establishing vital normative protections against gender-based violence, it lacks enforcement power. Its focus on long-term "economic empowerment" overlooks the immediate crisis of debt bondage, providing "too little, too late" for workers in recruitment debt.
- 'Safe and Fair' Initiative (EU & UN, 2019): This program integrates protection into National Action Plans via technical assistance. However, its impact is limited by the host country's prioritization; if legal frameworks in service-providing nations remain weak, abuse persists. Furthermore, it fails to penetrate the entire employer-industry chain.

3.1.2 National Government (Taiwan)

Despite MOHW guidelines, structural gaps persist between mandate and capacity:

- Reactive vs. Proactive Supervision: Current policies focus on 24-hour mandatory notification post incident. This reactive measure fails to identify "pre-event" warning signs due to manpower crises and a lack of routine inspections.
- Inaccessibility in Isolated Environments: Reliance on the 1955 hotline (approx. 550 cases/year) assumes communication autonomy. It provides no intervention for victims whose phones are confiscated or movements restricted.
- • Non-standardized Interpretation: While MOHW mandates interpreters, it lacks standardized sourcing or training. Deficiencies in legal terminology and trauma-informed care often led to poor evidence collection and secondary victimization.

3.1.3 NGOs and Advocacy Gaps

- IDWF: Faced with a legal ceiling, as many host countries forbid migrant unions, stripping workers of collective bargaining power (WIEGO & IDWF, 2019).
- HRW: The Advocacy-Action Gap; international pressure often takes years to influence policy, failing to provide immediate safety for victims currently in crisis (Hafner-Burton, 2008).
- MFA: Resource Fragmentation; transnational rescue chains break if local partners lack funding or face political pressure (ILO, 2017).
- GAATW: The Complexity Trap; focuses on academic "root causes" and human trafficking definitions can be too slow for grassroots crises, causing victims of "standard" abuse to fall through the cracks (GAATW, 2007).

Table 1: Gaps and Levers of Change

Challenge	Gap	Levers of Change
Systemic Power Imbalance	• The treaty is difficult to implement because it cannot be forced, and not many countries have realized the importance of this problem. • The bind between employers and female migrant workers is hard to break.	• Reformation of law: Female migrant workers should be entitled the right to switch employers at their will. • Transparent Intermediary System: Break the information asymmetry.
Legislative Loopholes and Administrative Negligence	• Local agencies violate procedural justice through legal shortcuts and inaction, causing delays that remain systemically ignored.	• Governments should partner with NGOs to resolve staffing shortages, ensuring legal frameworks are successfully implemented. • Define clear penalties for misconduct and procedural skips to close loopholes and ensure strict enforcement.
Global and Local Racial Hierarchies	• Perpetrators operate under the "protective umbrella" of systemic impunity. • Official statistics often fail to reflect the true scale of the crisis.	• Implement stricter legislation and introduce harsher penalties. • Digital Advocacy & Awareness Campaigns: Humanizing Labor. • Inclusion of Voices: Engaging Youth and Migrant Leaders.
Environment of Impunity	• Lax social supervision and major legal loopholes reinforce the predatory behaviour of perpetrators and further weaken the deterrent power of the law. • Physical isolation further creates "blind spots" for information and data. • Structural inequality gives employers an absolute power advantage, transforming the labour relationship into one of oppression. • Migrant workers choose to endure abuse because they are burdened by debt.	• Mandatory Digital Connectivity: Legislation should prohibit employers from withholding mobile phones and ensure workers have unmonitored internet access in private spaces. • Establishing "Cyber-Safe Zones" and Peer-to-Peer Alert Systems: Government and NGOs can establish migrant worker communities or "blacklists", so workers know which employers or locations are dangerous. • Financial Support: Establish a "mutual aid fund" or introduce microloans to help migrant workers repay high-interest loans.
Human Rights Fallacy	• Public Prejudice Widespread bias and lack of awareness normalize discrimination. Without accurate understanding, the public remains indifferent to the systemic struggles migrant workers face. • A critical fallacy treats basic rights as "extra benefits." When fundamental dignity is seen as an optional privilege, it undermines the legal and moral obligation to protect workers.	• Human rights education must transcend traditional textbook instruction. Documentaries and artistic media serve as more potent catalysts for change by humanizing systemic issues through narrative storytelling. These mediums effectively bridge the empathy gap, transforming abstract concepts into relatable lived experiences that challenge public prejudice.
Loopholes of Intermediary System	• Distorted agency evaluation scores allow non-compliant companies to maintain good ratings. • For brokers, the cost of not reporting is low, so they prefer to operate in legal gray areas or illegally force settlements.	• Strict Accountability: Lower the threshold for business suspension (e.g., reducing the number of permitted non-reporting violations) to make the law stricter. • Increase penalties for failure to report.

4 Conclusions

This study employs systems thinking to deconstruct the multi-layered structural violence faced by female migrant workers in Taiwan. The analysis reveals that sexual violence is not merely a collection of isolated criminal acts, but a systemic outcome driven by the intersection of physical isolation, economic debt bondage, and socio cultural dehumanization. The self-reinforcing loops of institutional apathy and information asymmetry create a "legal vacuum" where existing protections fail to reach those in private domestic spaces.

To dismantle the systemic exploitation of migrant workers, a multi-pronged approach is required. Policymakers should shift from reactive notification to proactive inspections and ensure digital connectivity to

bridge the "visibility gap." Legal practitioners must implement trauma-informed interpretation services to prevent secondary victimization and ensure judicial accessibility. Finally, intermediaries and NGOsshould reform the brokerage system to eliminate recruitment fees, thereby reducing the financial debt that currently compels victims to endure abuse in silence.

By addressing these systemic leverage points, Taiwan can move beyond symbolic compliance toward a robust, gender-sensitive protection framework that upholds the fundamental human rights of all migrant workers.

5 AI Declaration

We used Gemini during the development of this paper. The tool was employed to assist in summarizing academic literature, refining English sentence structures for clarity.

6 Ethical Statement

This study is based on a secondary analysis of publicly available judgments from the Taiwan High Court and reports from government and non-governmental organizations. We did not conduct any direct interviews for this study.

References

American Institute in Taiwan. (2024). Trafficking in persons report in Taiwan. Available at: https://www.ait.org.tw/2024-trafficking-in-persons-report-taiwan/

American Institute in Taiwan. (2025). Trafficking in persons report in Taiwan. Available at: https://www.ait.org.tw/2025-trafficking-in-persons-report-taiwan/

Amnesty International. (2021). Living in the shadows: The human rights of migrants. Anderson, B. (2000). Doing the dirty work? The global politics of domestic labour. Zed Books. Anti-Trafficking Review. (2016). 'View of villains and victims, but no workers: Why a prosecution-focussed approach to human trafficking fails trafficked persons', Anti-Trafficking Review, (6).

Bajracharya, R., & Sijapati, B. (2012). 'The Kafala system and its implications for Nepali domestic workers', Policy Brief, 1.

CEDAW. (2008a). General recommendation No. 26 on women migrant workers (CEDAW/C/2009/WP.1/R). United Nations. Available at:

https://asiapacific.unwomen.org/sites/default/files/Field%20Office%20ESEAsia/Docs/Publications/2013/CEDAW%20GR26%20on%20Women%20Migrant%20Worker.pdf

CEDAW. (2008b). General suggestions: female migrant workers.

Criminal Judgment of the Taiwan High Court. (2022). Taiwan High Court Criminal Judgment No. 258 of 2022 (Infringement Appeal Case).

Criminal Judgment of the Taiwan High Court. (2024a). Taiwan High Court Criminal Judgment No. 72 of 2024 (Infringement Appeal Case).

Criminal Judgment of the Taiwan High Court. (2024b). Taiwan High Court Criminal Judgment No. 107 of 2024 (Infringement Appeal Case).

European Union External Action. (2019). Safe & Fair: Women migrant workers project in ASEAN. Available at: https://www.eeas.europa.eu/node/64689_en

Global Alliance Against Traffic in Women (GAATW). (2007). Collateral Damage: The Impact of Anti-Trafficking Measures on Human Rights around the World. GAATW.

Hafner-Burton, E. M. (2008). 'Sticks and stones: Naming and shaming the human rights enforcement problem', International Organization, 62(4), pp. 689–716.

International Labour Organization (ILO). (2011). Domestic Workers Convention, 2011 (No. 189). International Labour Organization (ILO). (2017). Access to justice for migrant workers in South-East Asia. ILO Publishing.

International Labour Organization (ILO). (2024). Access to labour justice for migrant workers through cross border litigation.

Kapp, K. W. (1950). The Social Costs of Private Enterprise. Harvard University Press.

Lan, P. C. (2006). Global Cinderellas: Migrant domestics and newly rich employers in Taiwan. Duke University Press.

Linn, N., Kaewboonchoo, O., & Nanthamongkolchai, S. (2024). 'Violence against women and its effects on mental health and quality of life: A study of Myanmar migrant workers in Central Thailand', Journal of Migration and Health, 10, 100231. Available at: https://doi.org/10.1016/j.jmh.2024.100231 Migration Data Portal. (2025). Migration and humanity mobility: key global figures.

Ministry of Health and Welfare. (2024). Cross-tabulation of nationalities and occupations of foreign victims in sexual assault case reports. Available at: https://dep.mohw.gov.tw/DOPs/cp-1287-14966-105.html Ministry of Health and Welfare. (2025). Strengthening the Reporting Mechanisms for Sexual Assault Cases Involving Foreign Nationals and the Principles for Inter-agency Coordination and Task Allocation. Available at: https://dep.mohw.gov.tw/DOPs/cp-1287-14966-105.html

Ministry of Labor. (2025). Statistical Analysis of Complaints via the 1955 Labor Consultation and Complaint Hotline. Available at: https://statdb.mol.gov.tw/statiscla/webMain.aspx

Parreñas, R. S. (2015). Servants of globalization: Women, migration and domestic work (2nd ed.). Stanford University Press.

Rothna Begum / Human Rights Watch. (2015). “I Was Sold”: Abuse and Exploitation of Migrant Domestic Workers in Oman. Available at: https://www.hrw.org/report/2016/07/13/i-was-sold/abuse-and exploitation-migrant-domestic-workers-oman

Sassen, S. (1991). The Global City: New York, London, Tokyo. Princeton University Press. Simon, H. A. (1955). 'A behavioral model of rational choice', The Quarterly Journal of Economics, 69(1), pp. 99– 114.

SPA. (2024). Survey report on the status of gender-based violence against female migrant workers in the workplace. Available at: https://www.spa.org.tw/article/405

Taiwan Taipei District Court. (2024). Criminal Judgment No. 113-Qin-Su-Zi-107.

Thaler, R. H. (1999). 'Mental accounting matters', Journal of Behavioral Decision Making, 12(3), pp. 183–206. The Control Yuan. (2022). Investigation report 0007. Available at:

https://www.cy.gov.tw/CyBsBoxContent.aspx?n=133&s=17805

The Control Yuan. (2025). Investigation report 0018. Available at:

https://www.cy.gov.tw/CyBsBoxContent.aspx?s=49256

UN General Assembly. (2023). Violence against women migrant workers (Resolution 78/180). Available at: https://docs.un.org/en/A/80/317

UN Women. (2025). What is unpaid care work and how does it power the economy? Available at: https://www.unwomen.org/en/news-stories/explainer/2025/10/what-is-unpaid-care-work U.S. Department of State. (2024). 2024 country reports on human rights practices: Taiwan. Available at: https://www.state.gov/reports/2023-country-reports-on-human-rights-practices/taiwan/ WIEGO, & International Domestic Workers Federation (IDWF). (2019). Your toolkit on ILO Convention 189 - The Domestic Workers’ Convention. WIEGO.

Workforce Development Agency, Ministry of Labor. (2021). Three conditions for manpower agencies charging service fees to migrant workers. Available at:

https://www.wda.gov.tw/News_Content.aspx?n=31&s=3936

Zeng, Y. T. (2025). 'How could a caregiver be allowed to hire migrant workers after committing sexual assault?!', Right Plus. Available at: https://rightplus.org/2025/09/04/sexual-assault-worker/

Modern Dystopia: Social Media and Digital Fragility

Huang, Yu-En, Wu, Mei-Ling, Cheng, Hsiang-Wei
Baling High School, Taipei, Taiwan

Abstract

This practical paper investigates how social media platforms produce digital fragility through reinforcing relationships between platform incentives, algorithmic recommendations, high-stimulation interface design, and user behavior. Rather than treating excessive social media use only as a problem of individual discipline or content moderation, the paper frames it as a system-level governance challenge. The analysis focuses especially on adolescents, because peer evaluation, social belonging, and online feedback can make engagement-driven mechanisms more behaviorally influential during this developmental stage.

The paper applies a systems thinking approach to examine how social media platforms convert attention into economic value. Engagement generates behavioral data, behavioral data improves recommendation precision, and more precise recommendations increase further engagement. This circular process is strengthened by mechanisms such as infinite scroll, autoplay, push notifications, visibility metrics, and continuous recommendation feeds. To analyze these dynamics, the paper uses three systems tools: a Systems Map, an Iceberg Model, and a Stakeholders' Map. These tools show that problematic social media use is not produced by a single cause, but by the interaction of monetization models, recommendation architecture, user vulnerability, advertiser incentives, regulatory limitations, and unequal control over data and visibility.

Based on this analysis, the paper proposes a practical attention-governance framework designed to reduce cumulative exposure to high-stimulation mechanisms without regulating speech or content. The framework contains four pillars: a dual-layer protection mechanism across platform accounts and operating systems; transparent thresholds for activating protection; graduated reduction of high-stimulation mechanisms; and enforcement measures with anti-circumvention safeguards. Its core principle is to regulate mechanisms rather than content, shifting intervention from individual self-control to system design while avoiding centralized surveillance.

The paper also identifies unresolved limitations, including unauthenticated browsing, browser-level enforcement, cross-device and cross-platform circumvention, uncertainty around exact threshold values, risks of overbroad feature classification, possible misuse as content regulation, increased power for operating-system providers, and jurisdictional constraints. Overall, the paper contributes a system-informed practical framework for addressing social-media-driven digital fragility by targeting the feedback structures that sustain attention exploitation.

Keywords: System Dynamics, Social Media, Adolescents, Policy Intervention, Attention Exploitation
Paper type: Practical Paper

1 Introduction

Harms associated with excessive social media use are often framed as problems of individual discipline, digital literacy, or content moderation. This paper argues that such framing is incomplete. Social media platforms are not merely communication tools, but socio-technical systems in which business incentives, recommendation algorithms, interface design, user behavior, and regulatory limitations interact over time. For this reason, social media-related harm should be understood not only as an individual behavioral issue, but also as a system-level governance challenge. In this paper, digital fragility refers to a condition in which users' attention, autonomy, social judgment, and public discourse become increasingly vulnerable to platform systems that optimize for engagement, prediction, and retention.

In contemporary social media systems, user attention has become a measurable and economically valuable resource. Platforms are incentivized to increase engagement, extend time spent on the platform, improve advertising performance, and collect behavioral data. Recommendation systems support this process by ranking and amplifying content according to its likelihood of generating interaction. As users engage more frequently, platforms collect more behavioral data; this data improves recommendation precision; improved recommendations then increase further engagement. This creates a reinforcing feedback loop in which attention is repeatedly converted into data, and data into stronger attention capture (Zuboff, 2019; Casey, 2022; Milli et al., 2025).

In this paper, digital fragility refers to a condition in which users' attention, autonomy, social judgment, and public discourse become increasingly vulnerable to platform systems that optimize for engagement, prediction, and retention.

This issue is especially significant for adolescents. During adolescence, peer evaluation, social belonging, and social feedback often become particularly salient. Platform features such as likes, comments, follower counts, recommendation feeds, autoplay, push notifications, and infinite scrolling may therefore have a stronger behavioral influence on adolescents than on many adult users. The problem is not simply that adolescents lack

self-control, but that their developmental sensitivity can interact with platform mechanisms designed to encourage repeated use and prolonged engagement (Sherman et al., 2016; da Silva Pinho et al., 2024).

This framing also matters because many existing responses focus on symptoms rather than system structure. Education, parental guidance, screen-time reminders, and content moderation may be useful, but they do not fully address the incentives that make attention capture profitable and scalable. If the underlying system continues to reward engagement maximization, individual-level interventions may reduce some harm without changing the conditions that repeatedly reproduce it. A more effective response, therefore, requires attention to feedback loops, platform incentives, interface mechanisms, and the distribution of power between platforms, users, and regulators.

This paper adopts a systems thinking approach to examine these dynamics. It uses a Systems Map, an Iceberg Model, and a Stakeholders' Map to analyze how platform incentives, user responses, recommendation systems, data collection, and governance gaps form recurring feedback structures. Based on this analysis, the paper proposes a practical attention-governance framework that regulates high-stimulation mechanisms rather than content. It aims to reduce cumulative exposure to features such as autoplay, infinite scroll, and continuous recommendation feeds without censoring speech or creating a centralized surveillance system.

The objective of this paper is to develop a system-informed practical framework for reducing social-media-driven digital fragility by identifying the feedback structures that sustain attention capture and by proposing mechanism-based interventions that reduce cumulative exposure to high-stimulation platform features without regulating content.

2 Problem Framing

The problem examined in this paper is not simply that users spend too much time on social media. It is that social media platforms are built within incentive structures that reward repeated engagement, behavioral prediction, and long-term retention. From the platform perspective, attention is not only a sign of user interest but also a measurable resource that supports advertising exposure, data collection, and commercial value. The longer users remain active, the more opportunities platforms have to collect behavioral data, refine recommendation systems, and strengthen their position in the attention economy (Zuboff, 2019; Casey, 2022; Meta Platforms, Inc., 2025). This means that problematic use should not be understood only as an outcome of individual choice. It is also shaped by a business environment in which platforms compete to capture, retain, and monetize user attention. The term "high-stimulation mechanisms" is used here to describe interface and recommendation features that continuously supply new cues, reduce stopping points, or encourage repeated interaction, including infinite scroll, autoplay, push notifications, visibility metrics, and continuous recommendation feeds.

Recommendation systems are central to this structure. They do not merely show content in a neutral order. Instead, they rank, filter, and amplify content according to predicted interaction, such as clicks, likes, comments, shares, viewing time, and repeated visits. This means that content which is emotionally stimulating, rapidly consumable, or socially rewarding may receive greater visibility than content that is less interactive. Over time, such systems can create a feedback process in which engagement produces more data, data improves prediction, and improved prediction increases further engagement (Milli et al., 2025). Interface mechanisms such as infinite scroll, autoplay, push notifications, and continuous recommendation feeds reinforce this process by reducing natural stopping points and making repeated use easier to sustain (Rixen et al., 2023). Even when users intend to stop, the surrounding design environment may continue to provide new cues, rewards, and opportunities for re-engagement.

This platform-side dynamic interacts with user-side vulnerability. Social media is valuable because it provides communication, entertainment, information, and social connections. However, these ordinary uses can become entangled with mechanisms of social validation. Likes, comments, follower counts, and visibility metrics make social feedback immediate and measurable, linking social validation to repeated checking and potentially addictive patterns of use (Jain and Saleem, 2025). For general users, this can encourage repeated checking and prolonged participation. For adolescents, the influence may be stronger because adolescence is a period in which peer evaluation, social belonging, and social feedback often become especially important. Research on adolescent responses to social media feedback suggests that peer endorsement can influence attention, evaluation, and behavior in online environments (Sherman et al., 2016; da Silva Pinho et al., 2024). Therefore, adolescent vulnerability should not be understood as a personal weakness. Rather, it should be understood as a developmental sensitivity that may be intensified by engagement-driven platform design.

The historical development of social media has strengthened these dynamics. Earlier social media platforms were more closely associated with connecting friends, communities, and existing social networks. Content was

often presented through simpler timelines, and participation was more closely tied to known social relationships. As platforms expanded alongside smartphones, social ties became harder to leave, and engagement indicators such as likes, shares, and followers became more central to participation. In the algorithmic era, platforms increasingly shifted from chronological display toward personalized ranking and recommendation. This shift made social media less like a passive communication channel and more like an adaptive environment that continuously learns from user behavior (Nieborg and Poell, 2018; Twenge et al., 2019). As a result, the system became better at predicting what users are likely to watch, click, share, and return to.

The issue is also shaped by unequal power. Users participate in social media systems, but they do not meaningfully control the rules of visibility, ranking architecture, or data governance. Platforms possess greater access to behavioral data and greater ability to modify interface design, recommendation logic, and monetization strategies. This creates a form of data asymmetry: platforms can observe and optimize user behavior at scale, while users usually see only the final interface and have limited knowledge of how their attention is being organized. Regulators, by contrast, often face jurisdictional limits when dealing with global platforms that operate across devices, markets, and legal systems. This creates a governance gap: users experience the consequences of platform design, while platforms retain substantial control over the mechanisms that organize attention and visibility (Chan and Kwok, 2022; Königs, 2024).

For these reasons, excessive or problematic social media use should not be framed only as a failure of self-control. It emerges from the interaction of platform incentives, algorithmic recommendation, high-stimulation design, user vulnerability, competitive pressure, and weak governance capacity. This framing is important because it changes the policy question. The central issue is not only how to educate users to make better choices, but how to redesign or regulate the system conditions that repeatedly make attention capture profitable, scalable, and difficult to resist.

3 System Thinking Approach

This paper adopts a systems thinking approach because social media-related harm cannot be adequately explained through a single cause. Excessive or problematic use does not arise only from weak self-control, harmful content, platform design, or adolescent vulnerability in isolation. Rather, it emerges from the interaction of these factors over time. Systems thinking is useful in this context because it focuses on recurring patterns, feedback loops, structural incentives, and the assumptions that allow a system to reproduce similar outcomes. Instead of asking only what immediate event caused a problem, this approach asks how different parts of the system reinforce one another and why the same patterns continue to appear (Meadows, 1999; Sterman, 2001).

The analysis was conducted as a conceptual systems analysis rather than an empirical study. First, relevant literature on engagement-based ranking, adolescent social feedback, interface design, and platform governance was reviewed to identify recurring mechanisms. Second, these mechanisms were organized into system variables and actor categories. Third, causal relationships were mapped through reinforcing and balancing loops. Finally, the identified feedback structures were used to evaluate possible leverage points and develop a practical attention-governance framework.

The central system examined in this paper consists of platform companies, users, adolescents, advertisers, recommendation algorithms, interface mechanisms, operating system providers, regulators, families, educators, and civil society organizations. These actors and mechanisms do not operate independently. Platform companies design systems that maximise engagement and retention. Users respond to personalised content, social validation, and frictionless interface design. Advertisers benefit from attention and behavioral prediction. Regulators attempt to set rules, but often face jurisdictional and technical limits. Adolescents are not the only affected group, but they are especially important because social feedback and peer evaluation can be more influential during this developmental stage.

To analyze these relationships, the paper uses three systems tools: a Systems Map, an Iceberg Model and a Stakeholders' Map. The Systems Map visualizes the major variables and feedback relationships in the social media ecosystem. It shows how engagement, social validation, compulsive behavior, smartphone dependence, usage time, behavioral data and recommendation algorithms can form a reinforcing loop. In this loop, higher engagement produces more data; more data improves recommendation precision; improved recommendation increases personalized content; and more personalized content encourages further engagement. This structure helps explain why problematic use can persist even when individual users intend to reduce their screen time. The Systems Map also shows broader consequences, including polarisation, imitation, fake information and weakened meaningful public discourse. The Systems Map in Figure 1 visualizes the reinforcing relationships among engagement, behavioral data, recommendation precision and further engagement.

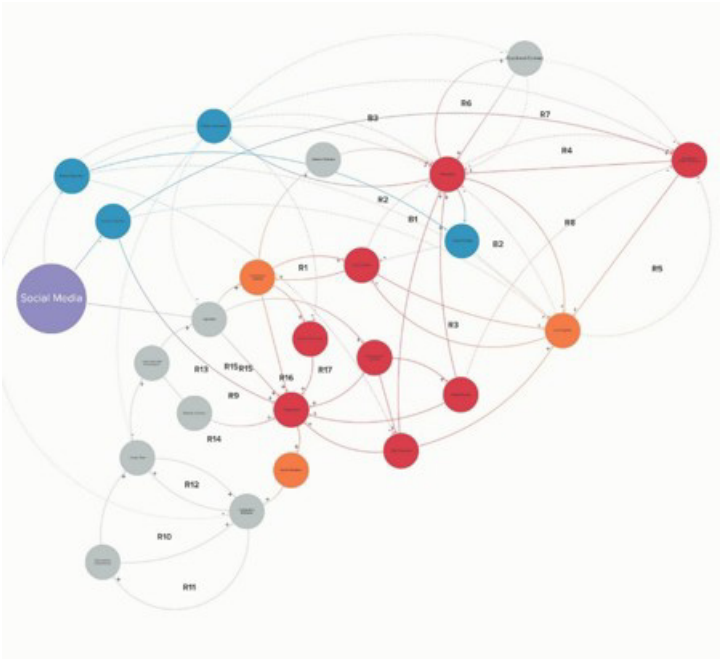


Figure 1: Systems Map
Source: Authors' Kumu systems map, available at: https://kumu.io/Ianlovesneko/cld#system-map

The Iceberg Model is used to separate visible symptoms from deeper system structures. At the event level, the issue appears as rising screen time, compulsive checking, increased phone dependence, and exposure to highly personalized content. At the pattern level, these events become recurring behaviors, such as repeated platform return, longer sessions, and growing dependence on recommendation feeds. At the structural level, the analysis identifies the mechanisms that produce those patterns, including engagement-based monetization, algorithmic ranking, infinite scroll, autoplay, push notifications, and feedback metrics. At the level of mental models, the system is supported by assumptions that constant connectivity, online visibility, and engagement growth are normal or desirable. This model is important because it prevents the analysis from stopping at surface-level symptoms. The Iceberg Model in Figure 2 separates visible symptoms from deeper patterns, structures, and mental models.

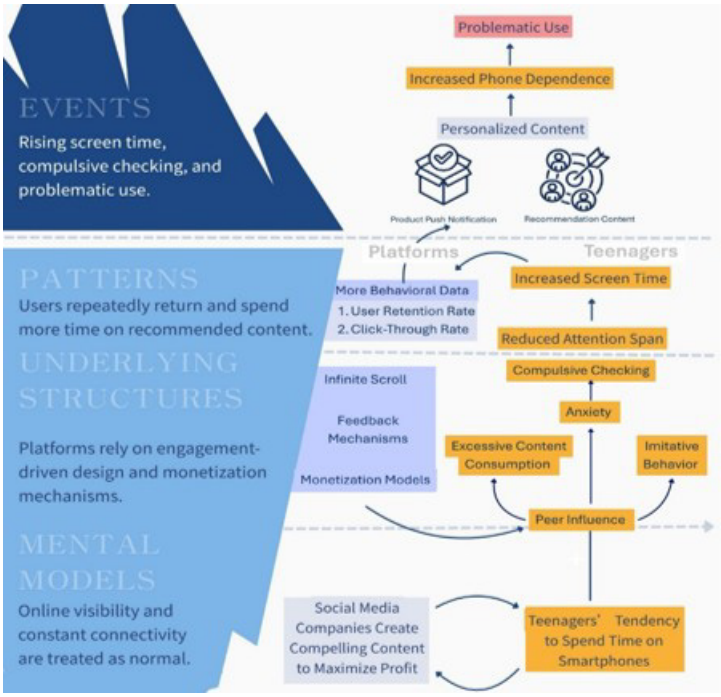


Figure 2: Iceberg Model

The Stakeholders' Map extends the analysis by identifying the actors and institutions that shape or are affected by the system. These include users, adolescents, families, educators, advertisers, platform companies, content creators, operating system providers, regulators, and civil society organizations. The map also recognises that some important forces are not human actors but structural mechanisms, such as recommendation systems, engagement metrics, and high-stimulation interface features. Treating these mechanisms as part of the stakeholder environment is useful because they influence behavior, distribute visibility, and shape decision-making even though they are not formal institutions. A full interactive version of the Stakeholders' Map is available at: https://kumu.io/Ianlovesneko/stakeholders-map

Together, these tools clarify why social media-related harm should be understood as a system-level governance problem. The Systems Map identifies reinforcing feedback loops, the Iceberg Model explains the deeper structures behind visible behavior, and the Stakeholders' Map shows the unequal distribution of influence among users, platforms, regulators, and technical infrastructures. This approach also supports the later policy analysis. If the problem is produced by reinforcing feedback structures, then effective intervention cannot rely only on education or individual discipline. It must also address the rules, information flows, and interface mechanisms that allow attention exploitation to become profitable and self-reinforcing.

4 Feedback Loops and Leverage Points

The systems thinking approach outlined in the previous section identifies that problematic social media use is sustained not by a single cause, but by recurring feedback structures. These feedback loops explain how platform incentives, user responses, algorithmic systems, and interface design interact to produce stable patterns of behavior over time. In this section, four major loops are identified: two reinforcing loops that intensify problematic use, and two balancing loops that may reduce or counteract these effects.

The first reinforcing loop (R1) can be described as the engagement loop. Higher levels of user engagement increase social validation through likes, comments, shares, and other feedback signals, which can encourage repeated checking and habit formation (Jain and Saleem, 2025). This encourages repeated checking and more frequent interaction, which can contribute to compulsive behavior and increased usage time. As usage increases, platforms collect more behavioral data, allowing recommendation systems to improve their precision. More precise recommendations then deliver more personalized and engaging content, which further increases engagement. This loop converts attention into data, and data into stronger attention capture, forming the core mechanism of the social media attention economy.

The second reinforcing loop (R2) connects personalized content, imitation, and polarisation. Recommendation systems tend to deliver content aligned with users' prior behavior and preferences. This can lead to repeated exposure to similar viewpoints, emotionally engaging material, and socially rewarded content. Such exposure may encourage imitation and strengthen belief consistency, contributing to polarisation and reducing the diversity of perspectives encountered. As users continue to interact with similar content, the system further reinforces the same pattern. Over time, this loop may weaken meaningful public discourse while increasing the visibility of highly engaging but potentially divisive material. (Avalle et al, 2024)

In contrast, two balancing loops represent mechanisms that can reduce or stabilize the system. The first balancing loop (B1) involves critical thinking and user awareness. Increased exposure to diverse perspectives, improved media literacy, and stronger critical evaluation can weaken the effects of polarisation and reduce the behavioral pull of high-stimulation content. However, this loop is relatively weak because it depends on user effort and education, while the reinforcing loops are continuously supported by platform design and economic incentives.

The second balancing loop (B2) involves external intervention, including platform-level adjustments or regulatory action. Measures such as limiting high-stimulation features, increasing transparency, or introducing friction can reduce excessive engagement and weaken the intensity of reinforcing loops. Unlike B1, this loop does not rely solely on individual behavior, but changes the conditions under which behavior occurs. However, current regulatory approaches often struggle to operate at the same scale and speed as platform-driven dynamics.

From a leverage point perspective, these loops highlight that not all interventions are equally effective. Parameter-level interventions, such as education and awareness campaigns, may support users but are unlikely to change the structural incentives that drive engagement. Stronger leverage points involve modifying information flows, feedback-loop intensity, and system rules. For example, increasing transparency in how recommendation systems operate can reduce information asymmetry. Reducing the intensity of high-stimulation mechanisms, such as autoplay and infinite scroll, can weaken reinforcing loops. Establishing

enforceable obligations and anti-circumvention rules can change the system conditions under which platforms operate.

Based on this analysis, the paper distinguishes between complementary interventions and core structural interventions. Complementary measures, such as digital literacy and critical thinking, play an important supporting role but are insufficient on their own. Core interventions must target the mechanisms that sustain reinforcing loops, particularly those related to engagement maximization, algorithmic amplification, and cross-platform attention capture. These insights provide the analytical foundation for the attention-governance framework proposed in the next section. From a systems perspective, effective interventions must target leverage points within reinforcing feedback structures rather than only addressing surface-level symptoms. Table 1 summarises the candidate interventions and their corresponding leverage levels.

Table 1: Candidate interventions by leverage level

Intervention	Target loop(s)	Meadows level	Role
Digital literacy and critical thinking	B1, R2	Supportive Information Flow	Complementary
Platform–OS feature-state reporting	R1	Level 6	Pillar 1
Threshold-triggered reduction of stimulation	R1, R2	Levels 7–8	Pillars 2–3
Legal duty and anti-circumvention	R1, B2	Level 5	Pillar 4

5 Proposed Attention-Governance Framework

Building on the feedback-loop analysis above, this paper proposes an attention-governance framework designed to reduce cumulative exposure to high-stimulation social media mechanisms. The framework does not aim to eliminate social media use, censor content, or create a centralized surveillance system. Its core principle is to regulate mechanisms rather than content.

In this paper, “high-stimulation mechanisms” refer to platform features that continuously supply algorithmically selected content, reduce natural stopping cues, or encourage repeated interaction without a deliberate user request. Examples include autoplay, infinite scroll, push notifications, and continuous recommendation feeds. These mechanisms matter because they shape how long users remain engaged, how easily they stop, and how often they return.

The first pillar is a dual-layer protection mechanism. This mechanism operates through two independent layers: a platform account layer and a device layer. The platform account layer addresses cumulative high-stimulation usage of the same account across multiple devices. The device layer addresses cross-platform usage on the same device and would be implemented through operating systems such as iOS or Android. Either layer can trigger protection independently when defined thresholds are reached. This design avoids creating a centralized cross-platform data system while still addressing platform-hopping behavior.

The second pillar is transparent and rule-based thresholds. Rather than relying only on opaque platform judgment, the framework requires clear trigger conditions for intervention, reflecting broader concerns about algorithmic transparency and ethical platform design (Hu, 2025). These may include cumulative high-stimulation usage time, continuous usage duration, late-night activity, and repeated switching between platforms. The purpose of these thresholds is not to punish users, but to define when the system should reduce stimulation. Users should also be able to understand why protection was activated and challenge incorrect triggers.

The third pillar is graduated mechanism regulation. Once thresholds are reached, platforms would progressively reduce high-stimulation features. Early intervention may involve reminders or pause screens. Stronger intervention may disable autoplay, interrupt infinite scrolling, reduce recommendation intensity, or switch users to chronological or subscription-based feeds, building on emerging research on contextual interventions during infinite scrolling (Meinhardt et al., 2025). The content itself remains accessible; what changes is the delivery mechanism. This distinction is essential because the framework aims to reduce attention exploitation without becoming content moderation or viewpoint control.

The fourth pillar is enforcement and anti-circumvention. Without enforceable duties, platforms may redesign systems to preserve engagement while appearing compliant. They may rename features, shift high-stimulation functions into alternative formats, or use notifications to pull users back into restricted environments. The framework, therefore, requires audits, compliance checks, and penalties strong enough to prevent circumvention from becoming economically rational. Protection-related data should also be used only for safety and regulatory purposes, not for advertising, commercial analysis, or recommendation optimization.

Together, these four pillars shift intervention from individual self-control to system design. Education and digital literacy remain important, but they are insufficient when the surrounding system is optimized to maximize engagement. A systems-informed policy must weaken the reinforcing loops that convert attention

into behavioral data and behavioral data into further attention capture. The proposed framework does this by changing information flows, reducing feedback-loop intensity, and creating enforceable rules around high-stimulation design.

6 Limitations and Future Refinement

This framework has several limitations that require further refinement before implementation. First, the dual-layer mechanism can reduce platform-hopping, but it cannot eliminate all forms of circumvention. Users may browse without logging in, move between browsers and apps, use multiple devices, or combine cross-device and cross-platform switching. Fully preventing these behaviors would require a centralized identity or data-sharing system, which this paper deliberately avoids because it would create serious privacy and surveillance risks. Therefore, the framework should be understood as a risk-reduction mechanism rather than a system of perfect control.

Second, exact threshold values remain uncertain. Usage time alone does not always indicate harm, because context, purpose, age, content type, and individual differences matter. For example, one user may spend time on social media for communication or learning, while another may be repeatedly exposed to high-stimulation recommendation feeds. The definition of "high-stimulation mechanisms" may also create borderline cases, such as news feeds, search results, streaming services, or educational platforms with recommendation functions. Future policy design should therefore include expert review, pilot testing, public consultation, and clear scope limitations.

Third, the framework creates governance risks. A mechanism-based system could be misused as content regulation if legal boundaries are not explicit. Device-layer enforcement may also increase the power of operating system providers such as Apple and Google, especially if they gain authority to classify or restrict platform functions. In addition, single-country implementation may face jurisdictional limits because major platforms operate across borders. These risks show that attention governance requires strong safeguards, including transparency, independent audits, appeal mechanisms, data minimization, proportional enforcement, and international coordination. These limitations do not invalidate the framework, but they show that implementation must be cautious, reviewable, and institutionally constrained.

7 Conclusion

This paper has argued that social-media-driven digital fragility should be understood as a system-level governance problem rather than only an individual behavioral issue. Excessive or problematic use does not emerge from user weakness alone. It is produced through the interaction of platform incentives, recommendation systems, high-stimulation interface design, user vulnerability, and limited regulatory capacity.

Using a systems thinking approach, the paper identified how engagement, behavioral data, recommendation precision, and further engagement form reinforcing feedback loops. It also showed that adolescents may be especially affected because online feedback, peer evaluation, and social visibility can become more influential during this developmental stage. These findings suggest that education and digital literacy remain necessary, but insufficient, if the surrounding system continues to optimize engagement and retention.

The proposed attention-governance framework responds to this problem by targeting mechanisms rather than content. Through dual-layer protection, transparent thresholds, graduated reduction of high-stimulation features, and anti-circumvention enforcement, the framework aims to reduce cumulative exposure without censoring speech or creating a centralized surveillance system.

The broader contribution of this paper is to connect systems analysis with practical policy design. By focusing on feedback structures and leverage points, it shifts the discussion from blaming users toward redesigning the conditions that shape digital behavior. Future work should test threshold design, clarify regulatory scope, and explore international coordination. Nevertheless, the central implication remains clear: meaningful intervention must address the structures that make attention exploitation scalable, profitable, and self-reinforcing.

8 Ethics declaration

This paper did not involve human participants, interviews, experiments, surveys, or the collection of personal data. The analysis was based on literature review, systems mapping, and publicly available information. Therefore, formal ethical clearance was not required.

9 AI declaration

The authors used Perplexity, ChatGPT, Claude, Microsoft Copilot, and Grammarly during the development of this paper. Perplexity was used to search for academic sources; all sources were opened, checked for DOI information and relevance to the topic, and manually reviewed before use. ChatGPT was used to translate human-written Chinese drafts into English, refine wording, generate critical questions, and support red-team debate. Claude was used for red-team critique, multi-model debate, and identifying weaknesses in the proposed framework. Microsoft Copilot was used for formatting support and layout suggestions, and Grammarly was used for grammar correction.

Some AI-assisted wording was incorporated into the paper after human review and revision. However, AI did not generate the core ideas, systems analysis, causal logic, systems maps, or policy framework. AI-generated suggestions were used only as support for language refinement, critical questioning, and presentation improvement. All final academic judgments and decisions on whether to adopt, revise, or reject AI-generated suggestions were made by the authors.

References

Avalle, M., Di Marco, N., Etta, G., Sangiorgio, E., Alipour, S., Bonetti, A., Alvisi, L., Scala, A., Baronchelli, A., Cinelli, M. and Quattrociocchi, W. (2024) 'Persistent interaction patterns across social media platforms and over time', Nature, 628(8008), pp. 582–589. https://doi.org/10.1038/s41586-024-07229-y

Casey, T. (2022) 'Rethinking engagement: Challenging the financial model of social media platforms', Harvard Data Science Review, 4(1). https://doi.org/10.1162/99608f92.f09b190b

Chan, N.K. and Kwok, C. (2022) 'The politics of platform power in surveillance capitalism: A comparative case study of ride-hailing platforms in China and the United States', Global Media and China, 7(2), pp. 131–150. https://doi.org/10.1177/20594364211046769

da Silva Pinho, A., Céspedes Izquierdo, V., Lindström, B. and van den Bos, W. (2024) 'Youths' sensitivity to social media feedback: A computational account', Science Advances, 10(43), eadp8775. https://doi.org/10.1126/sciadv.adp8775

Hu, Z. (2025) 'Research on the impact of social media algorithmic on user decision-making: Focus on algorithmic transparent and ethical design', Applied and Computational Engineering, 174, pp. 18–22. https://doi.org/10.54254/2755-2721/2025.PO24665

Jain, P. and Saleem, T. (2025) 'Hooked on Hearts: The role of social media validation in behavioral addiction', International Journal of Environmental Sciences, 11(24s), pp. 1342–1348. https://doi.org/10.64252/ddnkap76

Königs, P. (2024) 'In defense of "surveillance capitalism"', Philosophy & Technology, 37, Article 122. https://doi.org/10.1007/s13347-024-00804-1

Meadows, D.H. (1999) Leverage Points: Places to Intervene in a System. Hartland, VT: The Sustainability Institute. Available at: https://donellameadows.org/wp-content/userfiles/Leverage_Points.pdf

Meinhardt, L.-M., Elhaidary, M., Colley, M., Rietzler, M., Rixen, J.O., Purohit, A.K. and Rukzio, E. (2025) 'Scrolling in the Deep: Analysing contextual influences on intervention effectiveness during infinite scrolling on social media', arXiv preprint arXiv:2501.11814. https://doi.org/10.48550/arXiv.2501.11814

Meta Platforms, Inc. (2025) Annual Report / Form 10-K. U.S. Securities and Exchange Commission. Available at: https://www.sec.gov/Archives/edgar/data/1326801/000132680125000017/meta-20241231.htm

Milli, S., Carroll, M., Wang, Y., Pandey, S., Zhao, S. and Dragan, A.D. (2025) 'Engagement, user satisfaction, and the amplification of divisive content on social media', PNAS Nexus, 4(3), pgaf062. https://doi.org/10.1093/pnasnexus/pgaf062

Nieborg, D.B. and Poell, T. (2018) 'The platformization of cultural production: Theorizing the contingent cultural commodity', New Media & Society, 20(11), pp. 4275–4292. https://doi.org/10.1177/1461444818769694

Rixen, J.O., Meinhardt, L.-M., Glöckler, M., Ziegenbein, M.-L., Schlothauer, A., Colley, M., Rukzio, E. and Gugenheimer, J. (2023) 'The loop and reasons to break it: Investigating infinite scrolling behavior in social media applications and reasons to stop', Proceedings of the ACM on Human-Computer Interaction, 7(MHCI), Article 228. https://doi.org/10.1145/3604275

Shanmugasundaram, M. and Tamilarasu, A. (2023) 'The impact of digital technology, social media, and artificial intelligence on cognitive functions: A review', Frontiers in Cognition, 2, 1203077. https://doi.org/10.3389/fcogn.2023.1203077

Sherman, L.E., Payton, A.A., Hernandez, L.M., Greenfield, P.M. and Dapretto, M. (2016) 'The power of the like in adolescence: Effects of peer influence on neural and behavioral responses to social media', Psychological Science, 27(7), pp. 1027–1035. https://doi.org/10.1177/0956797616645673

Sterman, J.D. (2001) 'System dynamics modeling: Tools for learning in a complex world', California Management Review, 43(4), pp. 8–25. https://doi.org/10.2307/41166098

Twenge, J.M., Martin, G.N. and Spitzberg, B.H. (2019) 'Trends in U.S. adolescents' media use, 1976–2016: The rise of digital media, the decline of TV, and the (near) demise of print', Psychology of Popular Media Culture, 8(4), pp. 329–345. https://doi.org/10.1037/ppm0000203

Zuboff, S. (2019) The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power. New York: PublicAffairs.

When Wetlands Disappear: What Human Development Means for Nature?

Jimmy Mao, Leticia Hsu, Sophie Yang
Taipei Digital Experimental High School (T-School), Taipei City, Taiwan, R.O.C.

Abstract

Wetlands represent some of the most productive and vital ecosystems on Earth. Despite covering approximately 6% of the global land surface, these environments—which include marshes, fens, peatlands, and various water bodies—provide indispensable ecological services, including water filtration, flood mitigation, and significant carbon sequestration. Research indicates that wetlands store more than one-third of the world's soil carbon, making them critical buffers against climate change (Mitsch and Gosselink, 2015). However, these ecosystems are currently vanishing three times faster than global forests. This study utilizes a systems-thinking methodology to analyze the drivers of wetland loss, focusing on the systemic interactions among human development, economic investment, and environmental conservation, with specific reference to the Wugu Wetlands in Taiwan.

The degradation of global wetlands has reached a critical threshold. Since 1970, an estimated 22% of the world's wetlands have been lost, primarily due to human interventions such as drainage for agriculture, urban expansion, and industrial pollution (WWF, 2024). This phenomenon, often termed the "wetland squeeze," occurs when coastal and inland habitats are trapped between rising sea levels and encroaching human infrastructure. The humanitarian implications are profound, as more than one billion people worldwide rely on wetlands for their livelihoods, including fishing, rice cultivation, and tourism (Ramsar Convention Secretariat, 2007). In Taiwan, the Wugu Wetlands serve as a primary example of this tension, where historical land-use policies have prioritized urban development over ecological preservation.

This research employs a Sustainable System Development (SSD) framework to map the complex feedback loops governing wetland exploitation. The methodology follows the strategic management and business innovation techniques outlined by Yan (2022), focusing on the identification of "reinforcing" and "balancing" loops within a system. Data were synthesized from international bodies, including the Ramsar Convention and the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES), alongside local biological data from the National Park Service of Taiwan. To process complex global and local trends, the researchers utilized artificial intelligence models to assist in the analysis of large datasets and causal relationships. The scope of the study focuses on five systemic categories: the Core exploitation-conservation balance, the Economic Loop, Community Support, Impacts on Nature, and External Factors such as climate change and international policy.

Keywords: Wetlands, Biodiversity, Systems Thinking, Human Development, Conservation

Paper type: Academic Research Paper

1 Introduction

Wetlands are defined as Areas of marshes, lagoons, peatlands, intertidal zones, and water bodies that are natural or man-made. Despite covering just 6% of the Earth's land area, these "biological kidneys" function as the planet's primary natural filtration system(Convention on Wetlands, et al., 2025). By trapping sediment and neutralizing pollutants, they provide a level of water purification that human infrastructure can rarely replicate.

But we are currently facing a "wetland squeeze" where urban expansion, agricultural development, and rising sea levels are fragmenting and drowning these vital coastal and inland buffers. When a wetland is paved over or polluted, it doesn't just disappear; it releases portions of the world's largest soil carbon store (33% of the global soil carbon) back into the atmosphere, accelerating the very climate volatility they are meant to prevent.

The collapse of these systems isn't just an environmental statistic—it is a humanitarian crisis, as more than one billion people worldwide depend directly on wetlands for their livelihoods. (IPBES, 2019) This includes the fishing and tourism sectors in Taiwan, as well as the millions of rice farmers globally. For coastal cities, the loss of wetlands means that storm surges and floods, once manageable threats, have now become catastrophic events for millions of urban residents.

This essay highlights the urgent need to connect our understanding of wetland value—estimated at up to $39 trillion in annual benefits (Convention on Wetlands, et al., 2025)—with the alarming rate of their destruction. By exploring various articles and research, we aim to uncover the complex systems surrounding wetlands and examine the links between ecological health and human security. Ultimately, we aspire to envision a hopeful future for these essential ecosystems.

2 Methodology

2.1 Research Methods

With the help of Google, Google Scholar and AI models such as Gemini and ChatGPT to help us search for websites and data related to our topic, we collected data and research studies from Ramsar, Wetland International, Global Wetland Outlook (Convention on Wetlands, 2025), and many official local government websites. By researching these gathered data and studies, we analysed wetlands' exploitation all over the world.

2.2 Research Models

Through attending a two-month systems thinking course (MIT x SSD), we've created multiple different models according to our topic. One of the models that contributed the most to our research was the causal loop diagram. Reinforcing loop, a system structure where a change in a variable sets in motion a cycle of events that amplifies or compounds the original change. Balancing loop, a self-regulating, goal-seeking structure in systems thinking that counteracts change to maintain stability and equilibrium. By understanding the different connections between elements, we are able to define which key factors we want to further explore.

Current environmental literature widely recognizes wetlands as among the most biologically diverse and productive ecosystems on Earth. Often referred to as "biological supermarkets" or the "biological kidneys" of the planet, wetlands perform irreplaceable functions. The other model we've used to solidify our understanding of this topic is the iceberg model. The model allows us to look at the problem from different views and think of other solutions.

2.3 Research Scope

Due to geographical limitations, our research is primarily based on Wugu Wetlands in Taiwan. However, we've also implemented the data given by other organizations and countries. Most of the research and studies used in this report are dated between 2022 and 2026; however, older studies remain relevant and useful despite being outdated.

3 Findings

3.1 Wetlands to the World

The definition of wetlands varies depending on what each country or organization deems important. We have compiled the different definitions of wetlands adopted by different countries and organizations (Yi-Yuan Su (蘇義淵), 2020) :

- According to the UNESCO (1971), wetlands are defined as: "areas of marsh, fen, peatland or water, whether natural or artificial, permanent or temporary, with water that is static or flowing, fresh, brackish or salt, including areas of marine water the depth of which at low tide does not exceed six metres."
- According to the EPA(Environmental Protection Agency, the United States)(2025) the definition of a wetlands are defined as: "areas where water covers the soil, or is present either at or near the surface of the soil all year or for varying periods of time during the year, including during the growing season."
- According to the Wetland Conservation Act (Ministry of the Interior, R.O.C., 2024), wetlands are defined as "areas such as marshes, lagoons, peatlands, intertidal zones, and water bodies that are natural or man-made, permanent or temporary, static or flowing, and fresh, salty, or brackish, including sea areas where the water depth does not exceed six meters at the lowest tide."

The difference between UNESCO (1971) and the definitions in the Wetland Conservation Act (Ministry of the Interior, R.O.C., 2024) is how specific the Ramsar convention defines wetlands. We believe the reason for these differences is due to the different mass of areas each organization is in charge of and what their ultimate protection goal is. For example, the Wetland Conservation Act (Ministry of the Interior, R.O.C., 2024) is more focused on the wetland's land while the EPA (2025) is more interested in protecting the growth of plants.

3.2 Ramsar Convention on Wetlands

3.2.1 Global Standards and Local Divergence

According to the Convention on Wetlands (2025), it emphasizes a shift from reactive to proactive wetland policies that prioritize conservation, restoration, and the integration of natural capital into decision-making.

Central to these policies is alignment with the Kunming-Montreal Global Biodiversity Framework (Convention on Biological Diversity, 2022), which targets the restoration of at least 30% of degraded ecosystems and the conservation of 30% of land and waters by 2030. The Outlook outlines four primary pathways for policy transformation:

- Integrating natural capital valuation to capture the true wealth of wetland services in economic planning.
- Recognizing wetlands as global public goods essential to the hydrological cycle
- Embedding wetlands in innovative financial solutions, such as blue bonds and debt-for-nature swaps
- Scaling up nature-based solutions by unlocking a mix of public and private finance while simultaneously reforming or eliminating harmful subsidies that drive degradation.

Although the Ramsar Convention on Wetlands is reasonable and solid, we believe that there should be more flexibility on the definition of wetlands, as each country's situation and environment are unique and can't be limited to a certain continent's situation. In contrast to the Ramsar Convention on Wetlands' policies (UNESCO, 1971), Taiwan, even though not eligible to sign the policy, has set its own rules for protecting wetlands.

3.2.2 Policies on Wetlands in Taiwan

According to the Ministry of the Interior, R.O.C. (2024), it prioritizes the "no net loss" of wetland areas and functions through a hierarchy of management strategies. Specifically, the Act mandates that any development project must first seek to avoid impacting wetlands; if impacts are unavoidable, they must be minimized, and any remaining loss must be compensated through restoration or creation elsewhere. There are four main points on the Wetland Conservation Act (Ministry of the Interior, R.O.C., 2024):

- Wetland Impact Assessment: The manufacturers that plan to develop projects significantly affecting a wetland of at least national importance are required to submit a wetland impact assessment and mitigation plan.
- Graded Management: Wetlands are classified into three levels- International, National, and Local- with the highest level of protection and stricter permit requirements applied to those of International and National significance.
- Permitted Wise Use: The policy encourages "wise use", which allows for sustainable activities that do not degrade the ecological character of the wetland, such as eco-tourism, environmental education, and traditional aquaculture or agriculture.
- Ecological Compensation: If a wetland is lost or degraded due to a project of major public interest, the developer must provide compensatory wetlands that are equal to or greater in area and ecological function.

Although policies in Taiwan are yet as detailed and complete as the Ramsar Convention on Wetlands, it still has its uses on limiting the development of wetlands. Though we agree with most of the policy's ideas, we believe that the main focus shouldn't be only on the land but rather also include the animals and plants living above it.

3.3 The Biodiversity of Wetlands

3.3.1 Current State of Wetland Biodiversity

In addition to wetlands being "the Earth's kidneys", they are the home to 40% of all global species and provide ecosystem services. However, wildlife populations in freshwater ecosystems have suffered a staggering 85% decline on average since 1970, according to the Living Planet Report (World Wildlife Fund., 2024).

3.3.2 Expected State of Wetland Biodiversity

To meet global biodiversity and climate goals, the Kunming-Montreal Global Biodiversity Framework(Convention on Biological Diversity, 2022) has established ambitious targets for 2030, requiring the effective restoration of at least 30% of all degraded ecosystems and the conservation of at least 30% of land, waters, and seas. More specifically, at least 123 million hectares of wetlands will be placed under restoration while also managing 428 million hectares within protected areas or other effective area-based conservation measures. Focusing on having a biodiverse environment, in our opinion, is better as it allows us to build a resilient ecosystem that can naturally adapt to changes rather than relying on constant human intervention.

3.4 The Loss of Wetlands

3.4.1 Loss Analysis on Wetlands

Through our research, we found that roughly 6% of the Earth's surface is covered by wetlands, and approximately 35% of global terrestrial carbon is stored within them. According to the Global Wetland Outlook(Convention on Wetlands, et al., 2025), wetlands are disappearing three times faster than forests. Wetlands act as massive carbon sinks, storing organic matter in their soil and plant communities, which is critical for global climate regulation. Since 1970, about 22% of the world's wetlands have been lost due to human activities like drainage, pollution, and over-exploitation. This loss is occurring at an average rate of 0.52% per year, with fresh water ecosystems like lakes and inland marshes experiencing the most significant proportional declines.

3.5 Our Observations

Initially, we believed that measuring wetland loss by area was sufficient. However, through extensive review we found that the biodiversity of wetlands was more important and more valid. Biodiversity in freshwater wetlands plays a key role in maintaining the stability of their habitat functions. (Song et al., 2024)The biodiversity of wetlands serves as a more comprehensive indicator of the ecosystem health than mere physical size. While acreage tells us how much land remains, biodiversity reveals how well it functions.

Ultimately, we believe the global community must elevate the protection of wetland biodiversity from a niche ecological concern to a top priority. Through our research, we have come to see these environments not just as land, but as the Earth's biological infrastructure. We are concerned that focusing only on wetland size may overlook a silent crisis—the collapse of ecosystems.

4 Discussion

4.1 The Disappearing Wetlands

Since the 1700s, the expansion of human exploitation has severely damaged many natural ecosystems. Wetlands are no exception.

According to the Convention on Wetlands, et al. (2025), the Earth has lost about 1,837 million hectares of wetlands since the 1970s, which is roughly 22% of the wetlands that once existed.

However, available data in Taiwan focuses on the number of National Important Wetlands rather than their actual area, resulting in limited insight into wetland conditions.

4.2 System of Wetlands

We analyzed the system behind Wugu Wetland (New Taipei, Taiwan) and the related issues to clarify the challenges and problems within the overall wetland system. Then, we group the elements into five categories:

- The Core: the center of the system, and the critical part of the wetland system.
- The Economic Loop: a key feedback loop related to wetland exploitation.
- The Community Support: the role and actions of the community surrounding the wetland.
- Impacts on Nature: the significance of the wetlands for the natural environment.
- External factors: external factors such as government agencies, companies, and environmental organizations.

4.2.1 The Core

The core of the wetland system is formed by "Wetland Exploitation," "Wetland Area," and "Wetland Conservation". The entire system is closely linked to these three important elements.

Wetland exploitation reduces wetland area; conversely, if conservation efforts are made, wetlands can gradually recover and return to their original state.

In our case, wetland exploitation refers to constructing buildings, developing parks, or directly destroying wetlands to satisfy human needs.

Wetland conservation includes actions such as protecting the wetland environment and its wildlife, as well as implementing ecological compensation measures (Mei-Li Hsue, Hsing-Juh Lin, 2015).

4.2.2 The Economic Loop

In addition, we identified some elements and three reinforcing feedback loops linked to the "Wetland Exploitation" element, which we call the "Economic Loop". This loop is driven by economic activity, and its outcome is a change in the "Wetland Area". This destructive loop caused the wetland area to decrease rapidly.

The following statements are going to explain the details of the elements in the economic loop.

- **Company Investment:**
  Some companies invest in the exploitation of wetlands because they see them as business opportunities. As a result, they often promote wetland exploitation, which indirectly leads to wetland destruction—for example, through the construction of factories, parks, or other buildings.
- **Community Approval of Exploitation:**
  The communities surrounding wetlands sometimes approve wetland exploitation, and their level of support directly affects the extent of exploitation. Conversely, if the community doesn't support companies' development activities and prefers to preserve the environment, wetland exploitation may slow down.
- **Economic Growth:**
  Wetland exploitation can bring more economic activities and contribute to economic growth, but it also encourages further exploitation of wetlands.
  Economic growth in turn attracts more companies to invest in wetlands because of the potential financial returns.

At the same time, the perceived benefits of economic growth shape community attitudes toward wetland exploitation, making communities more likely to approve it.

The relationship among these elements creates three reinforcing feedback loops in the Economic System. These loops function as the main drivers of wetland exploitation and cause it to accelerate rapidly.

4.2.3 The Community Support

In addition to the core and the economic loop, community support is another component. The surrounding community also plays an important role in the system. Community support can have both positive and negative impacts on wetlands.

- **Community Environmental Knowledge:**
  Environmental knowledge strongly shapes the actions and attitudes of residents. With lower environmental knowledge, residents are more likely to approve of activities or policies that may lead to higher wetland exploitation. By contrast, higher environmental knowledge leads to greater support for wetland conservation.
- **Community Participation:**
  Higher levels of community participation tend to increase the willingness of residents to engage in conservation activities that would benefit the natural environment.
- **Community Identity:**
  Community identity is an important driver of community participation; residents would participate more if they think they are a part of the community.
- **Community Living Quality:**
  Living quality is a crucial concern for residents. Better living quality allows residents to live healthier and more securely. We believe that living quality is an important factor shaping community decision-making and that economic growth can improve living quality.

The community plays a key role in the system, reflecting public opinions and influencing decisions on wetland exploitation or conservation.

4.2.4 Impacts on Nature

Beyond economic activities and community support, it's important to discuss the roles of wetlands as part of the natural system.

- Wetlands provide habitats for many wildlife species. According to the Convention on Wetlands, et al. (2025), wetlands support about 40% of known species on Earth. In Taiwan, wetlands are often called "a paradise for birds," which highlights the importance of wetlands for animals, plants, and other creatures. In Taiwan, there are about 5,022 types of animals, plants, or other organisms that live in wetlands, including waterbirds, aquatic animals, aquatic plants, insects, and more. (National Park Service, MOI, R.O.C, 2026)

- Biodiversity helps maintain a stable food chain, which in turn affects the food available to humans. The plants, animals, and the environment itself form a complex web that may directly or indirectly influence human food resources. Wetland species and populations matter a lot for sustaining biodiversity.
- Wetlands have a water-purification function, which can reduce water pollution. However, water pollution caused by human activities can also destroy wetlands and hinder the water-purification function of wetlands.
- Wetlands play a regulatory role in the climate, helping to mitigate global climate change. Wetlands can modulate a high rate of greenhouse gases in the atmosphere, reducing about 60% of CH4(methane), 20% of CO2(carbon dioxide), and 6% of NO(nitrous oxide). (IPCC, 2007)

4.2.5 External Factors

Last but not least, some external factors are also taken into account to, to some extent, influence the system, such as:

- Policies:
  Some policies support companies in investing in or exploiting wetlands because governments aim to boost the economy of communities near wetlands.
- Support from non-governmental organizations (NGOs):
  Many NGOs that care about nature visit wetlands to carry out conservation work to help the wetlands recover.
  For example, in Taiwan, some NGOs have been contributing their efforts to protecting wetlands, such as the Society of Wilderness (社團法人中華民國荒野保護協會), Wetlands Taiwan (社團法人臺灣濕地保育聯盟) and the Taiwan Wetland Society (社團法人臺灣濕地學會).
- Government support:
  Some governments support NGOs or other wetland conservation initiatives in order to implement environmentally friendly policies. Since wetlands and riverside parks are publicly owned, NGOs must obtain authorization from the government before carrying out conservation activities.
- Company investment:
  Many companies invest in wetland exploitation or other forms of development because of the high land value. However, these activities can seriously damage the environment and may even completely destroy local ecosystems.
  For example, in Taiwan, Uni-President Enterprises Corporation has collaborated with the Society of Wilderness. Donation boxes have been placed in some 7-11 stores across Taiwan, not only to raise public awareness but also to support the Society of Wilderness in environmental conservation.

These factors are not part of the community or nature, but they can influence many decisions and serve as key drivers in the system.

The system map on the next page shows the overall system we identified in this issue, including the economic loop, community support, impacts on nature, and external factors.

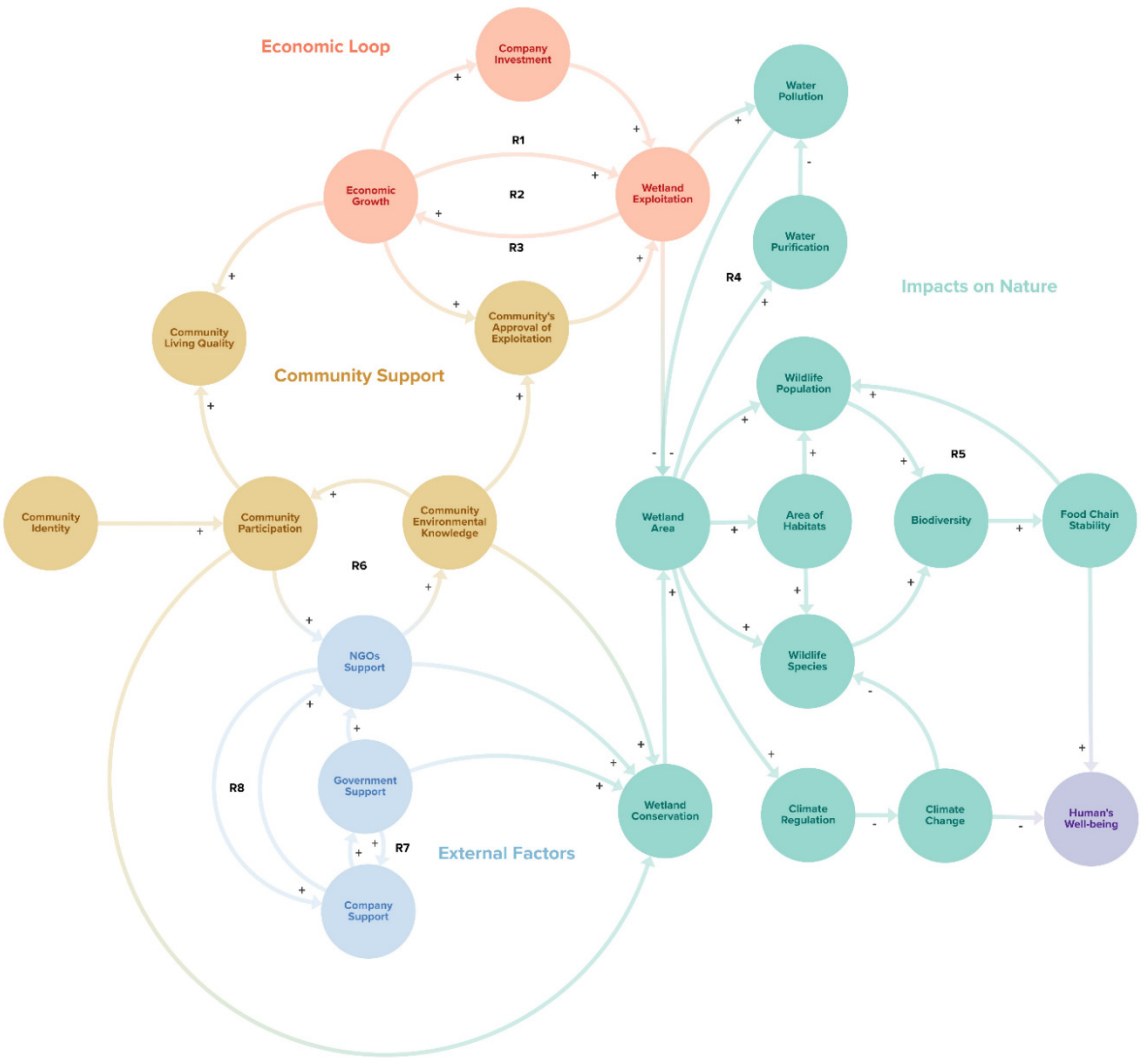


Figure 1 System Map of Wetlands with External Factors

5 Conclusion and Future Directions

5.1 Challenges

The over-development of wetlands must be addressed, but there are many complex challenges we need to confront. If we fail to tackle the following challenges, it will be very difficult to reduce wetland exploitation because of these terrible loops.

5.1.1 The Economic Loop and Community Expectation

For investors, wetlands are simply undeveloped land. If they develop the wetland, they gain financial returns, reinforcing an economic loop.

The community surrounding wetlands also benefits from economic growth and improvements in living quality, so they have little reason to oppose exploitation projects.

Governments can also gain from these exploitation projects – through taxes, economic feedback, or even votes – so policies often end up favoring investors, encouraging further wetland exploitation.

5.1.2 Climate Change

Climate change is partly due to wetland loss, but it also further accelerates wetland degradation. As wetlands disappear, we lose a crucial natural system for regulating the global climate and slowing global warming.

In addition, climate change harms human livelihoods as well as animals and plants around the world. People around the globe are increasingly concerned about climate change, which is driven not only by CO_2 emissions but also by the loss and degradation of natural ecosystems such as wetlands.

5.1.3 Water Pollution

Similarly, wetlands play a key role in purifying water quality, yet water pollution is also a major driver of ecological collapse within wetlands. As more wetlands are degraded, we not only weaken their water-purification capacity but also create a vicious cycle in which polluted water further damages the remaining wetlands and accelerates their loss.

The water pollution is also harmful for the animals, plants, and also humans. Sometimes, water pollution could cause severe diseases and even more serious ecological disasters, or ecological catastrophes.

5.1.4 Lack of Environmental Knowledge

The environmental knowledge of residents in the community is extremely important. If the community lacks sufficient environmental knowledge or awareness, residents may become the ones who damage nature—for example, by littering or feeding wild animals—and they will also be more likely to approve destructive actions such as wetland exploitation.

By contrast, if the community has sufficient environmental knowledge, residents can become protectors of the wetland. They are more likely to think critically about whether to approve development projects near wetlands, and they can participate in a wide range of wetland conservation activities that help improve and restore these ecosystems.

5.2 Suggestions

5.2.1 Implementing Natural Capital Accounting to classify wetlands practically.

We believe that governmental policies are crucial to wetland exploitation. Therefore, if we need to see changes in wetlands, close examination of policies and a shift in the current power dynamic are required.

Historically, there are countless development-oriented policies that exploit wetlands because the market value of a building was easily quantified, while the ecological services of a wetland remained "invisible" on a balance sheet.

To address over-development, Taiwan pioneered in Asia with the Wetland Conservation Act (implemented in 2015). It already classifies wetlands into "International," "National," and "Regional", three levels of importance. But to make classifying wetlands more detailed and practical, Natural Capital Accounting (NCA) should be implemented. This framework mandates that the economic value of flood mitigation, water purification, and carbon sequestration be integrated into provincial and national budgets. When the "shadow cost" of replacing a wetland with a multi-million dollar water treatment plant or sea wall is factored in, development projects that once seemed profitable are revealed as long-term economic liabilities.

By ending subsidies for coastal encroachment and prioritizing Nature-based Solutions (NbS), policies can align economic incentives with ecosystem preservation, ensuring that the "Core" wetland area is no longer a trade-off for growth, but a foundation for it.

5.2.2Implement "Citizen Science" for Community Support

To address the "Community Expectation" for quality of life, we must move beyond passive education to active community stewardship. Currently, people only care about wetlands when their own living quality has improved. We want to change it; the wetland shouldn't be a reward for a high quality of life; it should be the **engine** that produces it.

Instead of the community benefiting from the "Economic Loop" of exploitation, we must establish a system where **Wetland Health** directly drives **Local Well-being**.

We want to use citizen Science tools to gamify bio-monitoring: Citizen Science tools allow the public to voluntarily engage in scientific research. We can launch a "Community Health Dashboard," where residents can use mobile apps to log sightings of local species and assess water clarity. Through these tools, residents can access real-time data on how the wetland's "biological kidneys" are filtering their local water, cooling temperatures, and helping prevent expensive flood damage to homes. When people *understand* that a thriving wetland reduces their risks and expenses, supporting the community becomes a natural habit rooted in self-interest. And we will connect this data to local property values or "Eco-Credits." This initiative aims to create a

clear link between a thriving ecosystem and personal financial benefits, and also help residents recognize their role as stewards of the wetlands.

6 Ethics Declaration

Ethical Clearance was not required for this research.

7 AI Declaration

AI tools were employed solely for language refinement purposes and data gathering.

References

Convention on Biological Diversity. (2022). Kunming-Montreal Global Biodiversity Framework. https://www.cbd.int/gbf

Convention on Wetlands, Robertson, H., Fennessy, S., Hilton, G., Job, N., Kumar, R., Simpson, M., Aggestam, F., Aldred, M., Chacón, A., Costanza, R., Davidson, N., Field, C., Finlayson, C. M., Gandra, F., Gillis, L. G., Hernández-Blanco, M., Moritsch, M., Thornton, S., … Van 'T Hoff, V. (2025). Global Wetland Outlook 2025: Valuing, conserving, restoring and financing wetlands (First Edition). Convention on Wetlands. https://doi.org/10.69556/GWO-2025-eng

Environmental Protection Agency(EPA), United States. (2025). What is a wetland?.Wetlands. https://www.epa.gov/wetlands/what-wetland

IPBES. (2019). Global assessment report on biodiversity and ecosystem services. IPBES secretariat, Bonn, Germany.

IPCC (2007). Climate Change 2007 Synthesis Report. Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. https://doi.org/10.1017/CBO9781107415324.004.

Jeng-Rong Duh(杜政榮)(2005). The Sustainable Management of Wetland Environment of Taiwan(台灣濕地環境之永續管理). Journal of Living Sciences(生活科學學報). 9, 93~114 (in Chinese).

Lee, A., Lyu, A., Chang, A.-Y., Cheng, C.-H., Peng, C.-J., Hung, C.-H., Chiang, C.-Y., En-Li, C., Shiu, H.-J., Hsu, C.-T., Chang, H., Chen, H.-C., Lin, H.-S., Chen, I.-C., Ko, J. C.-J., Wang, J.-F., Lyu, J.-J., Tsai, J.-S., Huang, J. H.-W., … Yeh, Y.-M. (2020). The State of Taiwan's Birds 2020.

Lin, D., Chiu, C.-C., & Pursner, S. (2024). The State of Taiwan's Birds 2024.

MacKinnon, K. (n.d.). Other effective area-based conservation measures (OECMs) (PowerPoint slides). Convention on Biological Diversity.

Mei-Li Hsueh (薛美莉) & Hsing-Juh Lin (林幸助) (2015). Wetland Compensatory Mitigation: Mechanisms and Implementation in Taiwan (臺灣濕地生態補償之執行與推動). Journal of Wetlands (濕地學刊), 4(1), 22-30. (in Chinese).

Min-Ren Yan(顏敏仁)(2022). Handbook of Sustainable System Development for Strategic Management and Business Innovation(SSD企業創新策略管理技術). Future Career Publishing (前程文化).

Ministry of the Interior, R.O.C. (2024).Wetland Conservation Act (English Version).

Mitsch, W. J., & Gosselink, J. G. (2015). Wetlands. John Wiley & Sons.

National Park Service, MOI, R.O.C (2026). Biological Data. Taiwan Wetland. https://wetland-db.nps.gov.tw/#/bioDataPage

Ramsar Convention Secretariat. (2007). What are wetlands? (Ramsar Information Sheet no. 1).

UNESCO(1917).Convention on Wetlands of International Importance especially as Waterfowl Habitat. http://www.swan.org.tw/docdir/ESC4NXQ72L.pdf

Wardman, K. (2016, September 17). Balancing loop basics. The Systems Thinker. https://thesystemsthinker.com/balancing-loop-basics/

World Wildlife Fund. (2024). Living Planet Report 2024: A system in peril. WWF International.

Yi-Yuan Su (蘇義淵) (2020). Review the Definition of Wetland in Taiwan's Wetland Conservation Act from Comparative Law Perspective (以比較法觀點檢討濕地保育法中濕地的定義). Providence Law Review (靜宜法學) Vol. 9, 1–57.

Song, A., Liang, S., Li, H., & Yan, B. (2024). Effects of biodiversity on functional stability of freshwater wetlands: a systematic review. Frontiers in Microbiology, 15, 1397683. https://doi.org/10.3389/fmicb.2024.1397683

Montessori Education as a Human-Centered Foundation for AI Competency Development

David Coles, Tzu-Ting Hsu, Min-Ren Yan
North Star Bilingual Montessori, Taipei, Taiwan
National Chengchi University, Taipei, Taiwan

Abstract

The rapid integration of artificial intelligence (AI) is transforming how knowledge is created, interpreted, and applied within organizations. As AI increasingly automates routine and analytical tasks, sustainable competitive advantage is shifting toward distinctly human competencies, including abstraction, problem framing, adaptive learning, ethical judgment, and collaborative intelligence. These competencies shape how individuals engage with AI systems and contribute to hybrid human–machine knowledge processes (Brynjolfsson and McAfee, 2014).

From a knowledge management perspective, this shift raises a critical question: what educational foundations systematically prepare individuals for AI-mediated environments? While current discussions of AI competency emphasize technical literacy, long-term effectiveness depends on deeper cognitive and socio-ethical capacities developed through early learning experiences.

This paper proposes the Montessori–AI Capability Alignment Framework (MACAF), a three-layer conceptual model explaining how Montessori pedagogy supports the development of AI-relevant human capability. Layer 1 identifies key developmental design mechanisms embedded within Montessori education, including concrete-to-abstract progression, extended self-directed work cycles, self-correcting materials, mixed-age communities, and the teacher's role as guide. Layer 2 explains the human capabilities emerging from these mechanisms, such as structural abstraction, problem ownership, iterative reasoning, metacognitive awareness, and collaborative intelligence. Layer 3 demonstrates how these capabilities translate into AI-era organizational competencies, including algorithmic interpretation, effective human–AI interaction, iterative refinement, responsible AI use, and coordination within hybrid knowledge networks.

To illustrate the practical plausibility of the framework, the paper references North Star Bilingual Montessori School as an institutional illustration of how these mechanisms may be operationalized. Students engage in structured, real-world learning activities involving planning, collaboration, ethical decision-making, and adaptive problem solving, reflecting the formation of capabilities relevant to AI-mediated contexts.

This study contributes to knowledge management by linking developmental pedagogy with organizational capability formation. It argues that AI competency is not primarily a technical outcome, but the applied expression of human-centered cognitive and ethical development cultivated through educational design.

Keywords: Montessori education; AI competency; knowledge management; human capability; future talent development

Paper type: Conceptual Academic Paper

1 Introduction

Artificial intelligence (AI) is rapidly transforming how knowledge is generated, interpreted, and applied within contemporary organizations. As AI systems increasingly automate routine, analytical, and decision-support processes, sustainable competitive advantage is shifting toward distinctly human capabilities. In AI-mediated environments, individuals are required not only to use technological tools, but also to frame problems, interpret outputs, adapt to evolving contexts, and collaborate within hybrid human–machine systems (Brynjolfsson and McAfee, 2014).

This transformation raises important questions within knowledge management. Existing research has extensively examined how organizations create, transfer, and leverage knowledge assets (Nonaka and Takeuchi, 1995). However, comparatively less attention has been given to the developmental origins of the cognitive and socio-behavioral capacities that enable knowledge creation in the first place. Capabilities such as abstraction, adaptive learning, ethical judgment, and collaborative sense-making are not acquired instantly at the point of workforce entry. Rather, they emerge through earlier developmental experiences and learning environments.

In parallel, educational frameworks designed for the digital age emphasize similar human-centered competencies. For example, the Sustainable System Design (SSD) initiative developed by National Chengchi University (NCCU), aligned with Taiwan's 12-Year Basic Education framework, positions lifelong learning as its central objective. It is structured around autonomous action, communication and interaction, and social

participation, further elaborated into competencies such as systems thinking, adaptive innovation, ethical awareness, and collaborative engagement (NCCU, n.d.).

This convergence suggests that many competencies required for AI-mediated work represent an extension of long-recognized developmental capacities. The relevant question therefore becomes not whether these competencies matter, but how they can be systematically cultivated. While contemporary AI competency discourse often emphasizes technical literacy, such approaches risk overlooking the cognitive and socio-ethical foundations that enable meaningful AI interaction.

This paper proposes the Montessori–AI Capability Alignment Framework (MACAF), a three-layer conceptual model explaining how Montessori education functions as a developmental foundation for AI-era competency. The framework conceptualizes capability formation as a structured process linking learning environment design mechanisms, the emergence of human capabilities, and their later translation into competencies relevant to AI-mediated organizational contexts.

To clarify terminology used throughout the paper, **capabilities** refer to underlying cognitive and behavioral capacities (e.g., metacognitive awareness, iterative reasoning) that develop over time. **Competencies** refer to applied performance capacities in organizational contexts (e.g., responsible AI usage, human–AI coordination). MACAF explains how educational design mechanisms generate capabilities, which later manifest as competencies in AI-integrated work environments.

By articulating this developmental pathway, the paper advances a mechanism-based explanation of capability formation that complements existing knowledge management and AI competency perspectives. North Star Bilingual Montessori School is referenced as an institutional illustration demonstrating how such mechanisms may operate within an authentic educational setting.

2 Literature Background

2.1 AI competency and human-centered requirements

The rapid adoption of AI has led to increasing attention on the competencies required for effective participation in AI-mediated environments. While early discussions of AI competence emphasized technical skills such as programming and data analysis, recent perspectives highlight the importance of human-centered capabilities. These include abstraction, systems thinking, problem framing, adaptive learning, ethical reasoning, and collaborative intelligence (Brynjolfsson and McAfee, 2014).

In practice, the effectiveness of AI systems depends significantly on human interaction. The quality of outputs generated by AI models is shaped by the clarity of task definition, the ability to interpret results, and the capacity to iteratively refine prompts, data, and decision criteria. AI competency therefore increasingly involves higher-order cognition and judgment rather than purely technical proficiency.

2.2 Knowledge management and capability formation

Knowledge management research has long examined how organizations create, share, and apply knowledge. Nonaka and Takeuchi (1995) conceptualize knowledge creation as a dynamic process involving interactions between tacit and explicit knowledge, emphasizing human cognition, experience, and interpretation.

However, while knowledge management literature explains how knowledge is generated within organizational settings, it often assumes the presence of cognitive and socio-behavioral capabilities required for knowledge creation. In AI-mediated contexts, where organizational knowledge processes increasingly involve interaction with intelligent systems, understanding how these foundational human capabilities develop becomes particularly important.

2.3 Educational frameworks in the digital age

Educational frameworks developed in response to digital transformation increasingly emphasize transferable competencies rather than isolated subject knowledge. The SSD framework developed by NCCU provides an example of this orientation. It emphasizes lifelong learning through three core dimensions: autonomous action, communication and interaction, and social participation. These are further elaborated into competencies including systems thinking, adaptive innovation, ethical awareness, and collaborative engagement (NCCU, n.d.).

Such frameworks highlight that AI-era competencies are rooted in broader developmental capacities rather than in technical training alone. This perspective suggests that educational design should be examined as a foundational driver of long-term capability formation.

2.4 Montessori education as a developmental theory of capability formation

Montessori education offers a distinctive developmental theory grounded in the formation of human capability through interaction with a prepared environment (Montessori, 1967). Rather than viewing learning as the transmission of knowledge from teacher to student, Montessori conceptualized education as a process of self-construction emerging through purposeful activity, social interaction, and engagement with carefully designed materials. Montessori described education as “the natural process carried out by the child” through interaction with the environment (Montessori, 1967). In this view, education is understood as a process through which the individual constructs himself through interaction with the environment (Montessori, 1949).

Several core principles underpin this developmental approach. The prepared environment provides structured opportunities for exploration and independent activity. Auto-education positions the learner as an active agent in their own development, where cognitive, behavioral, and social capacities emerge through structured engagement with materials and environment. Human tendencies, including exploration, order, communication, and abstraction, are viewed as natural drivers of learning. Freedom within limits balances autonomy with responsibility, while control of error enables learners to identify and correct mistakes independently. Montessori also emphasized education for adaptation, arguing that educational systems should prepare individuals to respond effectively to changing social conditions (Montessori, 1948).

From a theoretical perspective, Montessori education can be understood as a mechanism-based developmental system. Its structural features generate recurring patterns of cognitive, behavioral, and social engagement which gradually give rise to transferable human capabilities. Key mechanisms include concrete-to-abstract progression, extended self-directed work cycles, embedded feedback through self-correcting materials, mixed-age peer learning, and the teacher's role as observer and guide. Auto-education (a term used in Montessori literature to describe aspects of the child’s self-directed learning through interaction with the environment) refers to the developmental process through which the child actively constructs cognitive, behavioral, and social capacities through purposeful interaction with a prepared environment, rather than through direct transmission of knowledge (Montessori, 1967; 1949).

These mechanisms do not directly produce isolated competencies. Instead, they shape how learners engage with experience, supporting the gradual development of independence, abstraction, concentration, adaptability, and social responsibility across contexts.

2.5 Conceptual integration and framework positioning

The preceding discussion highlights convergence across AI competency discourse, knowledge management theory, and contemporary educational frameworks. AI competency literature emphasizes human-centered abilities required for effective AI use. Knowledge management literature underscores the centrality of human interpretation in knowledge creation. Educational frameworks similarly emphasize transferable competencies that support adaptation and systems thinking.

However, these perspectives are rarely integrated into a developmental explanation of capability formation. This paper contributes by articulating a conceptual framework that links educational design mechanisms to capability development and subsequent AI-era competency expression.

3 Conceptual Framework: Montessori–AI Capability Alignment Framework (MACAF)

3.1 Framework overview

The Montessori–AI Capability Alignment Framework (MACAF) is proposed as a conceptual model explaining how educational design contributes to the formation of capabilities relevant to AI-mediated environments. Rather than treating AI competency as a set of technical skills, the framework conceptualizes AI readiness as a developmental outcome shaped by long-term patterns of cognition, behavior, and social participation.

The framework consists of three layers:

1. Developmental design mechanisms
2. Human capability formation
3. AI-era organizational competencies

These layers represent a sequential developmental process: mechanisms embedded in the learning environment shape recurring patterns of engagement, which generate capabilities that later manifest as competencies in AI-mediated organizational contexts.

3.2 Layer 1: Developmental design mechanisms

The first layer focuses on structural design mechanisms derived from Montessori developmental theory. These mechanisms operationalize principles such as the prepared environment, Auto-education, freedom within limits, and control of error. Rather than emphasizing content transmission, they shape how learners engage with knowledge, experience, and social interaction.

A key mechanism is the progression from concrete experience to abstract representation. Learners interact with physical materials that embody logical and mathematical relationships before transitioning to symbolic reasoning. This process supports the internalization of structure, allowing learners to move between levels of representation.

Extended, uninterrupted work cycles provide conditions for sustained engagement. Learners are given time to plan, execute, and reflect on their work, fostering persistence and responsibility.

Self-correcting materials introduce immediate feedback without external judgment. Errors are treated as information guiding revision, normalizing iteration and strengthening adaptive learning processes.

Mixed-age environments create opportunities for peer-based interaction. Learners observe, imitate, and teach one another, supporting perspective-taking and social responsibility. The teacher functions as observer and guide, shifting agency toward the learner.

Collectively, these mechanisms function as a developmental system through which learners gradually construct independence, concentration, abstraction, and social responsibility. Rather than directly producing competencies, they create recurring patterns of engagement that support long-term capability formation.

3.3 Layer 2: Human capability formation

The second layer explains how Montessori mechanisms generate specific human capabilities through repeated interaction. These capabilities are transferable and not confined to specific subject matter.

Concrete-to-abstract progression supports **structural abstraction**. Learners develop the ability to recognize patterns and relationships and apply them across contexts.

Sustained self-directed work cultivates **problem ownership** and intrinsic motivation, enabling learners to initiate tasks, manage time, and persist without constant external prompting.

Embedded feedback supports **iterative reasoning**, as learners become accustomed to testing approaches, revising strategies, and improving outcomes.

The teacher-as-guide structure strengthens **metacognitive awareness**, encouraging learners to monitor their thinking and adjust their approach.

Mixed-age social participation fosters **collaborative intelligence** and ethical awareness through shared responsibility, negotiation, and reciprocal teaching.

3.4 Layer 3: Translation into AI-era competencies

The third layer explains how capabilities formed in Layer 2 become operational as competencies in AI-mediated organizational contexts.

Structural abstraction supports **systems literacy** and the ability to interpret algorithmic outputs.

Problem ownership translates into effective **human–AI interaction**, as meaningful AI output depends on accurate task framing and constraint definition.

Iterative reasoning aligns with refinement cycles characteristic of AI-supported workflows.

Metacognitive awareness supports **responsible AI usage**, including critical evaluation, bias detection, and reflective judgment.

Collaborative intelligence supports coordination within **hybrid human–AI knowledge networks**, where work increasingly involves distributed collaboration mediated by AI systems.

The capabilities identified in MACAF align closely with contemporary frameworks such as SSD, suggesting that Montessori mechanisms support competency domains recognized within digital-age education discourse (NCCU, n.d.).

This translation is consistent with the NCCU Sustainable System Design (SSD) AI competency framework, which organizes AI competency across human-centred mindset, ethics of AI, AI techniques and applications, and AI system design, progressing from understanding to application and creation (Figure 1).

AI competency framework for students

COMPETENCY ASPECTS	PROGRESSION LEVELS		
	UNDERSTAND	APPLY	CREATE
Human-centred mindset	Human agency	Human accountability	Citizenship in the era of AI
Ethics of AI	Embodied ethics	Safe and responsible use	Ethics by design
AI techniques and applications	AI foundations	Application skills	Creating AI tools
AI system design	Problem scoping	Architecture design	Iteration and feedback loops

National Chengchi University (NCCU) (Year) Sustainable System Design Framework.
Figure 1: AI competency framework for students (NCCU Sustainable System Design Framework).

From a Montessori perspective, these competencies can be understood as contemporary expressions of broader developmental capacities associated with adaptation, independence, and social participation. Montessori education was not designed to prepare learners for specific technical tasks, but to cultivate flexible human capacities capable of responding to changing social conditions. Montessori emphasized that education must prepare individuals "for the conditions of modern life" (Montessori, 1948). In AI-mediated environments, these capacities acquire renewed significance.

3.5 Mechanism-based interpretation

MACAF provides a mechanism-based interpretation of capability formation. Rather than assuming competencies can be directly taught, it emphasizes that competencies emerge from long-term developmental processes shaped by structured interaction with the environment.

This shifts attention away from short-term outcome measurement and toward process explanation. It suggests that the design of learning environments may represent a foundational driver of AI-era human capability.

4 Institutional Illustration: North Star Bilingual Montessori School

4.1 Institutional context

North Star Bilingual Montessori School provides an institutional illustration of how the mechanisms articulated in MACAF may be operationalized within an authentic educational setting. The school adopts Montessori pedagogy as its core educational approach while emphasizing long-term capability development rather than short-term performance outcomes.

The practices observed at North Star reflect core Montessori principles including the prepared environment, control of error, normalization through concentration, peer learning, freedom within limits, and indirect guidance. Rather than positioning AI as a discrete subject, the school cultivates foundational cognitive, behavioral, and social capacities that align closely with the demands of AI-mediated environments.

4.2 Operationalization of developmental mechanisms

The developmental design mechanisms identified in MACAF are observable in the daily learning environment of North Star Bilingual Montessori School. The examples presented in this section are intended as illustrative demonstrations of Montessori developmental principles in practice. They are not presented as empirical evidence of AI competency outcomes, but as examples of how the mechanisms identified in Layer 1 may contribute to the capability formation described in Layer 2.

Prepared learning environments are structured to support the progression from concrete experience to abstract understanding. Learners engage with materials that embody mathematical, linguistic, and logical relationships, allowing them to internalize structural patterns prior to symbolic representation.

Extended, uninterrupted work cycles are embedded into the school schedule, providing learners with sustained periods of self-directed activity. Within these cycles, students are responsible for selecting tasks, managing time, and completing work sequences, thereby reinforcing autonomy and task ownership. For example, North Star's morning work cycle typically runs from 8:30am to 11:30am. During this uninterrupted period, students begin by writing in their journals and planning their independent work, selecting activities of their choosing while also organizing their time around scheduled presentations with guides. Learners may work individually with Montessori materials, collaborate on longer-term group projects, prepare for going-out activities, tend to the garden, or engage in practical responsibilities that support the classroom community. This sustained structure provides learners with the time and psychological space to initiate, persist, and complete meaningful work without reliance on constant external direction.

Self-correcting materials and embedded feedback structures are integrated into classroom activities. Rather than relying on external evaluation, learners receive immediate information through interaction with the task itself, encouraging revision and independent problem solving. For example, when students at North Star use command cards for the checkerboard material to complete long multiplication, they are able to check their final answer by turning over the command card to reveal the correct result. If their answer is incorrect, the child can identify the discrepancy without being judged or corrected by an adult and is encouraged to revisit the process. In many cases, the learner may attempt the problem again, seek clarification from a classmate, or return to a more familiar material such as the stamp game or bead frame to isolate the source of error before retrying the checkerboard. This structure normalizes error as part of learning and supports iterative reasoning through repeated cycles of action, feedback, and refinement.

Mixed-age classroom organization facilitates peer-based learning and creates conditions for reciprocal teaching. At North Star, classrooms typically include children aged 6, 7, 8, and 9 years old, which naturally encourages younger learners to seek guidance from older, more experienced peers. In many cases, an older child's explanation of a concept or demonstration of a material can be clearer and more accessible than an adult presentation, reinforcing and deepening the younger child's understanding. At the same time, older children often take pride in supporting younger classmates, and the act of teaching helps consolidate their own learning by requiring them to articulate steps, clarify reasoning, and model appropriate work habits. Importantly, this dynamic is not strictly hierarchical, as younger children who master a skill early may also guide older peers, demonstrating that competence is not determined solely by age. Through observation, imitation, and explanation, learners engage in admiration-based modeling and gradually transition from follower to leader, strengthening perspective-taking, social responsibility, and collaborative intelligence within the learning community.

The teacher's role is primarily that of an observer and guide rather than a central instructor. Instead of directing learning continuously, guides monitor student engagement, diagnose points of difficulty, and intervene only when necessary to restore productive independence. For example, during a fraction equivalence activity, a child may repeatedly attempt to match fraction pieces incorrectly, becoming frustrated and beginning to avoid the task. Rather than immediately correcting the child or providing the answer, the guide observes carefully to identify where the misunderstanding occurs, such as confusing the size of units or misinterpreting the relationship between parts and the whole. The guide may then offer a brief, targeted prompt or micro-presentation, such as asking the child to compare two fraction pieces directly or to rebuild the whole using different combinations, before stepping back again. This approach maintains learner agency while strengthening metacognitive awareness, as the child is encouraged to reflect on their own reasoning process and adjust strategies independently.

Through these practices, the school operationalizes the developmental mechanisms forming the first layer of MACAF.

4.3 MACAF mapping table: mechanisms to competencies

The table below summarizes the alignment between Montessori mechanisms, the capabilities they generate, and the AI-era competencies they support, using North Star as an institutional illustration.

Montessori Principle	MACAF Mechanism	Capability Generated	AI-Era Competency	North Star Illustration
Prepared Environment	Concrete-to-abstract progression	Structural abstraction	Systems literacy and algorithmic interpretation	Use of Montessori materials that embody mathematical relationships before symbolic work

Montessori Principle	MACAF Mechanism	Capability Generated	AI-Era Competency	North Star Illustration
Freedom Within Limits	Extended uninterrupted work cycle	Problem ownership and self-regulation	Effective human–AI task framing and independent execution	8:30am–11:30am work cycle with journaling, planning, independent work, projects, and going-outs
Control of Error	Self-correcting materials	Iterative reasoning	Iterative refinement in AI-supported workflows	Checkerboard multiplication with command cards and answer verification
Human Tendencies / Social Learning	Mixed-age learning community	Collaborative intelligence and ethical awareness	Coordination in hybrid human–AI knowledge networks	Peer teaching across ages 6–9 through reciprocal explanation and modeling
Auto-education / Indirect Guidance	Teacher as observer and guide	Metacognitive awareness	Responsible AI use and critical evaluation	Fraction equivalence support through observation, minimal prompting, and targeted intervention

4.4 Emergence of human capabilities

The structured implementation of these mechanisms contributes to the emergence of the capabilities described in Layer 2. Students develop pattern recognition and relational thinking through concrete materials, supporting structural abstraction. The continuity of self-directed work strengthens problem ownership, intrinsic motivation, and self-regulation. Interaction with feedback-rich materials supports iterative reasoning and adaptive learning.

The teacher-as-guide structure reinforces reflective thinking, strengthening metacognitive awareness. Participation in mixed-age communities supports the development of collaborative intelligence, communication competence, and ethical sensitivity. These capabilities emerge through sustained engagement rather than being introduced as isolated learning objectives.

4.5 Alignment with AI competency frameworks

The capabilities cultivated within this institutional illustration align with broader competency frameworks for the digital age. In particular, they correspond closely to SSD competency domains, including systems thinking, adaptive innovation, ethical awareness, and social participation (NCCU, n.d.).

This alignment suggests that Montessori developmental mechanisms are not only theoretically relevant but also consistent with contemporary competency frameworks used to define AI readiness.

4.6 Illustration interpretation

North Star illustrates how the conceptual structure of MACAF can be instantiated in a real educational setting. The school does not treat AI as a discrete technical discipline. Instead, it cultivates the foundational cognitive and social capacities required for meaningful engagement with AI-mediated environments.

This illustration should not be interpreted as empirical validation, but as an applied example demonstrating plausibility and operational coherence.

5 Discussion

5.1 Theoretical implications

This paper contributes to knowledge management by extending AI competency discourse into a developmental explanation of capability formation. While AI literature emphasizes the growing importance of human-centered competencies, and knowledge management emphasizes the role of human interpretation in knowledge creation, these perspectives are rarely linked through an explicit developmental pathway.

MACAF provides such a pathway by positioning educational design as an early-stage foundation for organizational capability. In this view, AI readiness is not simply a product of workplace training, but the applied expression of long-term developmental capability formation.

5.2 Implications for educational design

The framework suggests that educational systems seeking to prepare learners for AI-mediated environments should prioritize the design of learning conditions that cultivate transferable human capabilities. Rather than treating AI readiness as tool-based training, educational institutions may benefit from structuring environments that support autonomy, iterative feedback, reflective thinking, and collaborative responsibility.

5.3 Implications for organizations and AI-era talent development

For organizations, MACAF suggests that effective AI integration depends on workforce capabilities that cannot be developed through short-term training alone. Organizations may therefore benefit from broadening talent development approaches to include capability-focused development models that emphasize cognition, ethics, and collaborative intelligence.

5.4 Further directions

MACAF provides a foundation for future empirical investigation. Future research may examine how Montessori-aligned mechanisms influence measurable capability outcomes over time, or compare Montessori environments with alternative pedagogical systems. Further work may also refine alignment between MACAF and established competency frameworks to support policy and curriculum design.

6 Conclusion

This paper has examined how educational design contributes to the formation of capabilities relevant to AI-mediated environments. As artificial intelligence increasingly reshapes organizational knowledge processes, human-centered capacities such as abstraction, problem framing, adaptive learning, ethical awareness, and collaborative intelligence become increasingly important.

The Montessori–AI Capability Alignment Framework (MACAF) contributes a developmental perspective to knowledge management by linking Montessori educational theory to capability formation and subsequent AI-era competency expression. Grounded in principles including the prepared environment, Auto-education, control of error, human tendencies, and education for adaptation, the framework explains how long-term patterns of interaction within educational environments may give rise to capabilities that later become operational within AI-mediated contexts.

The institutional illustration of North Star Bilingual Montessori School demonstrates how Montessori mechanisms can plausibly operate in practice to cultivate transferable capabilities aligned with AI-mediated organizational requirements. Overall, the study suggests that AI readiness is not primarily a technical outcome, but the applied expression of human cognitive and ethical development cultivated through educational design.

7 Ethics Declaration

This study is conceptual in nature and does not involve human participants, personal data, or experimental procedures. The institutional illustration referenced in this paper is used solely for explanatory purposes and does not include identifiable individual data. Therefore, no ethical approval was required.

8 AI Declaration

The authors acknowledge the use of AI-assisted tools (including language models) to support language refinement and editing. All conceptual development, analysis, and academic content were developed by the authors. The authors take full responsibility for the integrity and originality of the work.

References

Brynjolfsson, E. and McAfee, A. (2014) The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies. New York: W.W. Norton & Company.

Montessori, M. (1948) To Educate the Human Potential. Madras: Kalakshetra Publications.

Montessori, M. (1949) The Formation of Man. Madras: Theosophical Publishing House.

Montessori, M. (1967) The Absorbent Mind. New York: Holt, Rinehart and Winston.

Nonaka, I. and Takeuchi, H. (1995) The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation. New York: Oxford University Press.

National Chengchi University (NCCU) (n.d.) Sustainable System Design (SSD) Framework. Available at: https://www.canva.com/design/DAGznj91D5I/_IXvLjiachMQOsuew2kMRA/view?utm_content=DAGznj91D5I&utm_campaign=designshare&utm_medium=link2&utm_source=uniquelinks&utlId=ha2623416bc#8

The Relationship between ESG and Firm Efficiency with the Moderating Role of Digital Transformation and Knowledge Capability

Duong Minh Tran, Wen-Min Lu
Chinese Culture University, Taipei, Taiwan

Abstract

Given the complexities of contemporary global business environment, the incorporation of environmental, social, and governance (ESG) principles into business practices has become increasingly vital. At the same time, the adoption of digital technology has accelerated the effectiveness of ESG initiatives, which has drawn increasing scholarly attention. Several empirical studies have examined the association between ESG and digital transformation, whereas others have explored the link between ESG and corporate efficiency. Nevertheless, research on the moderating effect of digital transformation on the relationship between ESG performance and firm efficiency remains limited.

National knowledge capability can be considered as one of the main drivers of digital transformation. For instance, governments may enhance digital knowledge diffusion across society by allocating national budgets to educational infrastructures, such as schools and research institutions, or they can subsidise large-scale R&D projects related to digital technology. The national investment in education may enhance digital capability of labour force, and thereby fostering the effectiveness of digital transformation process of the whole organisation.

While prior scholarly research has analysed the relationship between firm-level knowledge capability and digital transformation, little is known about whether country-level knowledge capability moderates the impact of digital transformation on the association between ESG and corporate efficiency.

Grounded in trade-off and absorptive capacity theories, this study aims to address the above-mentioned research gaps by examining how digital transformation shapes the ESG – firm efficiency relationship, while incorporating national knowledge capability as a second-order moderator. Our empirical analysis draws on a dataset of 46 technology companies listed on NASDAQ from 2019 to 2023, obtained from the London Stock Exchange Group (LSEG) and the Lens database. For digital transformation measure, textual analysis is employed, extracting technology-specific vocabularies from firms' annual reports. To examine multiple dimensions of corporate efficiency, while tackling the shortcomings of conventional data envelopment analysis (DEA), we adopt Range Directional Measures Network DEA model (RDMDEA). The results show that integrating digital transformation to ESG may undermine short-term efficiency within economies characterized by high knowledge intensity. Nonetheless, the role of digital transformation requires the application of knowledge to be realised. Subsequently, absorptive capacity should be taken into consideration through the incorporation of digital transformation into ESG initiatives, to ensure that both knowledge and digital technology can generate commercial values.

Keywords: Digital Transformation; ESG; Knowledge Capability; Corporate Efficiency; Technology.

Paper type: Academic Research Paper

1 Introduction

In recent years, business operations have increasingly adopted environmental, social, and governance (ESG) practices to fulfil regulatory requirements and stakeholder expectations. ESG initiatives, viewed strategically, foster competitive advantages and organizational legitimacy, particularly under stricter regulatory regimes (Le and Lu, 2025). Integrating ESG into corporate strategy helps enterprises build stronger stakeholder ties, enhance reputation and brand recognition, and reduce environmental and social risk exposure (Kandpal et al., 2024; Fernández et al., 2015). These positive impacts, in turn, can result in greater operational efficiency and enduring value creation (Endrikat et al., 2021; Xie et al., 2019).

In spite of the above-mentioned benefits, the association between ESG performance and firm efficiency remains ambiguous, as existing literature reports mixed results. While some studies find positive relationships (Ahmad et al., 2021; Chang et al., 2023; Pham et al., 2022), others discover negative impacts (Cao et al., 2024), as well as non-significant effects (Wanke et al., 2023; Yu et al., 2021). Hence, the inconsistencies in existing literature underscore the role of additional factors in the association between ESG and corporate efficiency.

Among these factors, digital transformation denotes the integration of digital technologies into organizational processes, which fundamentally alters and modifies firms' operations and value creation. Digital technologies allow firms to optimize resource allocation and promote green innovation (Wang and Yang, 2024). Additionally, they can improve transparency, information processing, and support more effective decision-

making, thus accelerating ESG implementation. In this regard, digital transformation reinforces ESG initiatives and improve efficiency outcomes.

Nonetheless, digital transformation entails significant investment, which requires firms to consistently acquire, integrate, apply emerging knowledge characterized by uncertainty. Empirical evidence highlights that most digital transformation initiatives are undermined by insufficient knowledge and capabilities (Xiong, 2024). The success of implementation mainly depends on firms' capacity to acquire expertise and reconfigure organizational processes (Crupi et al., 2020). These requirements impose substantial pressures on firms' internal resources and learning capacity.

Knowledge plays a pivotal role, though its effects are not consistently beneficial. Successful firms not only acquire technological knowledge but also develop the absorptive capacity to apply it within organizational processes (Chanias et al., 2019). Absorptive capacity captures an enterprise's capability to recognize, integrate, and employ external knowledge (Cohen and Levinthal, 1990). A lack of absorptive capacity may hinder firms' ability to translate knowledge into performance gains. Matarazzo et al. (2021) reveal that deficient absorptive capacity constrains digital transformation, especially when employees lack the proficiency to use new technologies. Furthermore, incorporating new knowledge into established processes is costly and time-intensive, potentially diminishing short-term efficiency (Roberts et al., 2012; Xiong, 2024).

Ultimately, these points indicate that combining ESG initiatives with digital transformation might lead to burden instead of mutual benefit. Dual investment in ESG and digitalization often creates a 'capability overload,' where firms struggle to absorb, process and effectively implement both sets of changes at once.

Building on these premises, this research investigates the impact of digital transformation (DTI) on the ESG-efficiency nexus and explores how country-level knowledge capability (KC) conditions this relationship. Despite extensive study of ESG and digital transformation as individual drivers of performance, their combined impact—particularly across both financial and non-financial efficiency metrics—remains largely underexplored.

This study adopts a range directional measures network DEA (RDMDEA) model, using a dataset of 46 NASDAQ-listed technology companies over the period 2019–2023 to analyse operational, innovation, profit, and sustainability efficiency. Additionally, we adopt textual analysis method to collect DTI data from firms' annual reports, and integrate country-level KC as a contextual moderator. Truncated regression is employed to examine the association between ESG performance, DTI, KC, and firms' efficiency.

The rest of the paper is organized as follows. Relevant literature and hypotheses development are provided in section 2, while section 3 describes the data, variables, and methodology. Section 4 presents the empirical results, and section 5 covers the conclusion with directions for further research.

2 Literature Review

2.1 ESG and Firm Efficiency

This study integrates stakeholder and trade-off theories to examine the complex relationship between ESG performance and corporate efficiency. Stakeholder relationships can be enhanced by effective ESG initiatives, which subsequently improves brand image, and reinforces access to capital, ultimately bolstering overall firm performance (Freeman, 2010; Endrikat et al., 2021; Xie et al., 2019; Malik, 2015).

Nonetheless, investment in ESG may neither generate values immediately, nor secure corporate efficiency. According to the trade-off theory, ESG initiatives consume excessive financial and organizational resources, which may dilute managerial focus on fundamental operations (Preston and O'Bannon, 1997). In capital-intensive sectors, ESG investments — which require significant expenses — can become a burden on short-term profitability and operational efficiency (Patten, 2002; Hassel et al., 2005; Friedman, 2007). Therefore, we propose the following hypothesis:

H1: In short term, ESG performance is adversely associated with corporate efficiency.

2.2 Digital Transformation as a First-Order Moderator

The Stakeholder theory asserts that firms' responses to stakeholder expectations can be more effective through digital technologies, which enhance transparency and ultimately bolster ESG performance (Xu et al., 2022).

Nevertheless, digital transformation entails substantial investments in infrastructure, systems, and operational realignment. In the early stage of digital implementation, firms encounter elevated operating expenses characterized by considerable uncertainty (Gebauer et al., 2020; Han et al., 2014). In particular, the beneficial nature of digital transformation commonly exhibits a time-lag, meaning that firms may encounter

temporary performance degradation during the transition process (Kriebel and Debener, 2019; Shakina et al., 2021). Regarding the trade-off perspective, significant resource allocation for digital transformation may constrain firms' capacity to maintain ESG investments (Li et al., 2025). Thus, we propose the second hypothesis as follows:

H2: Digital transformation significantly impacts the relationship between ESG performance and corporate efficiency.

2.3 Knowledge Capability and Absorptive Constraints

Successful digital transformation is contingent upon firms' ability to acquire, integrate, and exploit knowledge effectively. This ability is defined by absorptive capacity theory as a firm's capability to identify, assimilate, and apply external knowledge to productive ends (Cohen and Levinthal, 1990).

The organisational learning process is significantly determined by national knowledge capability of the country in which a firm operates. Robust national innovation systems accelerate knowledge creation, diffusion, and skill development, which subsequently drive digital transformation (Qi et al., 2024; Ge et al., 2023). When operating in such environments, enterprises enjoy enhanced exposure to cutting-edge knowledge and technological resources, thus boosting innovation and efficiency.

However, national knowledge capability does not automatically translate into enhanced effectiveness for all firms. As digital technologies rapidly disrupt traditional processes, firms may struggle to absorb and integrate new knowledge into their operations, especially large-scale enterprises characterized by complex systems (Guo et al., 2023; Horváth and Szabó, 2019). Developing the required capabilities for keeping pace with technological change involves extensive investments in training, systems modification, and organisational restructuring, thereby escalating costs and hampering short-term efficiency (Brunetti et al., 2020).

Surplus knowledge acquisition may induce organizational complexity and conflict between emerging and existing processes (Feliciano-Cestero et al., 2023; Zhong and Ren, 2023), which in turn diminishes returns. Under such circumstances, the integration of digital technologies within knowledge-rich environments may exacerbate efficiency pressures related to ESG commitments.

H3: Knowledge capability amplifies the negative moderating impact of digital transformation on the relationship between ESG performance and corporate efficiency.

3 Research Design

3.1 Samples and Data Collection

We gathered a dataset of technology firms listed on the NASDAQ stock exchange from diverse sources, encompassing the London Stock Exchange Group (LSEG), Lens database, and corporate annual reports from 2019 to 2023. In particular, ESG scores and financial data were drawn from LSEG, while the patent-related data, including patent applications, granted patents, and design rights, were extracted from the Patent Lens platform. Digital transformation metric was drawn from firms' annual reports by using Python to collect the digital-related words frequency, and ultimately constructing the digital transformation index (DTI) by employing Principal Component Analysis (PCA). Knowledge capability metric, which is country-level data, was obtained from The World Intellectual Property Organisation (WIPO). Samples with missing data were removed to ensure the integrity and consistency. Following the data screening and cleaning procedures, the final sample consists of 46 companies with 230 firm-year observations. The research framework of this study is depicted in figure 1.

3.2 Research Framework

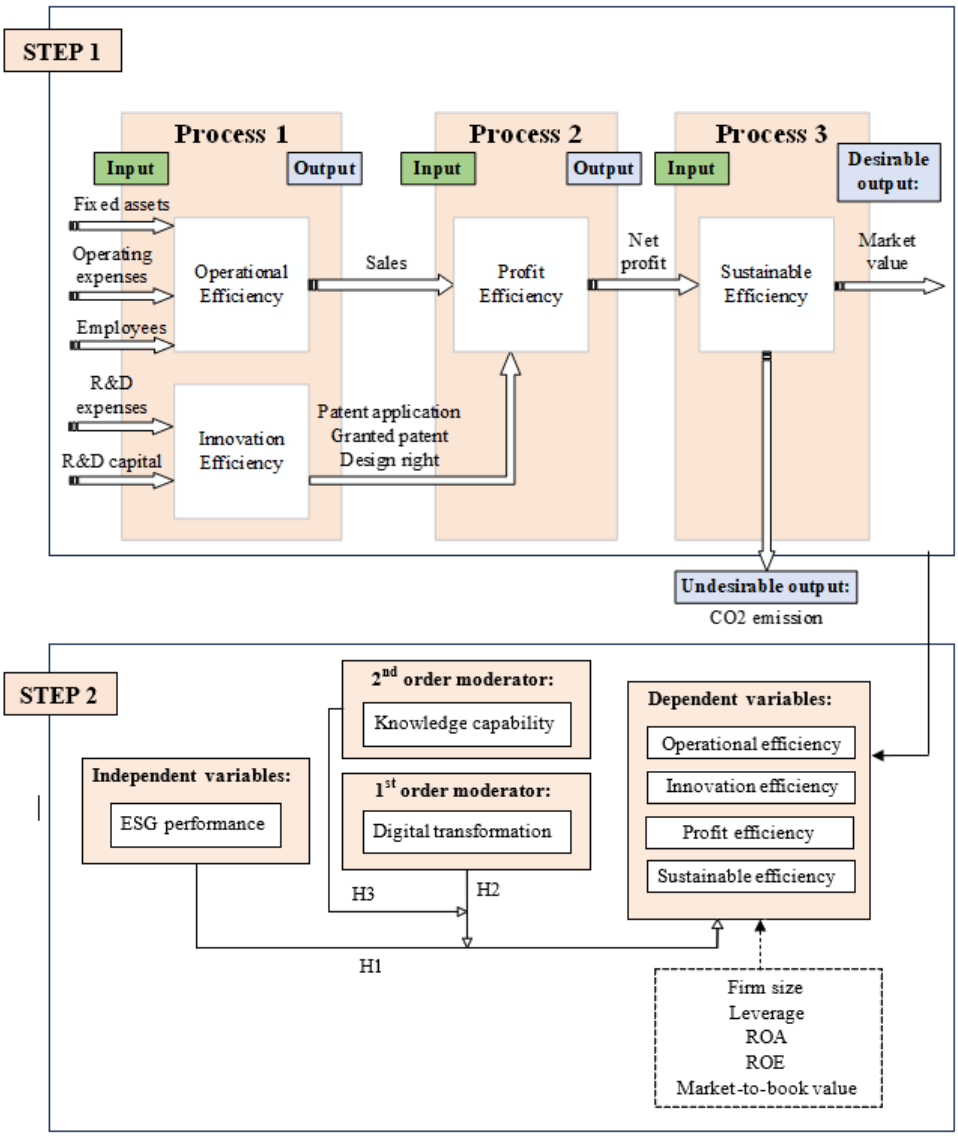


Figure 1 | Theoretical framework

4 Empirical Analysis

4.1 Descriptive Statistics Analysis:

Table 4 presents descriptive statistics and correlation matrices for the variables in RDMDEA model. We conducted the Kolmogorov–Smirnov test (K-S test), indicating that all inputs and outputs across the three processes follow nonnormal distributions with p-values below 0.01. This result suggests that inputs and outputs are valid for RDMDEA analysis since this method imposes no normality assumption.

Table 4 | Descriptive statistics and correlation matrices for RDMDEA analysis

Variables	Means	S.D.	1	2	3	4	5	6	7	8	9	10	11	12	K–S Test[a]
1. Fixed assets	10,192,577	25,425,616	1.000												$p < 0.01$
2. Operating expenses	18,937,515	45,651,856	0.79***	1.000											$p < 0.01$
3. Employees	28,743	42,616	0.87***	0.87***	1.000										$p < 0.01$

Variables	Means	S.D.	1	2	3	4	5	6	7	8	9	10	11	12	K–S Test[a]
4. R&D expenses	3,561,107	7,661,275	0.95***	0.88***	0.88***	1.000									*p* < 0.01
5. R&D capital	8,101,427	17,245,029	0.96***	0.88***	0.89***	0.90***	1.000								*p* < 0.01
6. Sales	26,579,269	65,196,288	0.79***	0.90***	0.88***	0.89***	0.89***	1.000							*p* < 0.01
7. Patent application	1,058	1,840	0.71***	0.60***	0.81***	0.69***	0.71***	0.69***	1.000						*p* < 0.01
8. Granted patent	545	955	0.78***	0.76***	0.87***	0.75***	0.77***	0.76***	0.89***	1.000					*p* < 0.01
9. Design right	21	84	0.53***	0.92***	0.65***	0.67***	0.67***	0.90***	0.53***	0.59***	1.000				*p* < 0.01
10. Net profit	6,431,386	17,257,044	0.78***	0.95***	0.88***	0.87***	0.87***	0.98***	0.68***	0.75***	0.83***	1.000			*p* < 0.01
11. Market value	159,924	434,945	0.61***	0.84***	0.8***	0.72***	0.72***	0.87***	0.60***	0.68***	0.74***	0.93***	1.000		*p* < 0.01
12. CO2 emission	849,027	1,794,005	0.86***	0.56***	0.76***	0.73***	0.74***	0.57***	0.68***	0.73***	0.29***	0.58***	0.43***	1.000	*p* < 0.01
***$p < 0.01$ [a]Kolmogorov–Smirnov test															

4.2 Hypothesis Testing

Table 7 | Correlation matrices for the main variables

Variable	1	2	3	4	5	6	7	8	9	10	11	12
1. ESG	1.00											
2. KC	0.14**	1.000										
3. DTI	0.08	-0.154**	1.000									
4. OPE	0.15**	-0.001	-0.024	1.000								
5. IE	0.28***	0.075	0.197***	0.048	1.000							
6. PE	-0.11	0.185***	-0.208***	0.133**	-0.229***	1.000						
7. SE	-0.17**	-0.218***	-0.153**	0.333***	-0.127	0.092	1.000					
8. LnFS	0.62***	0.018	0.326***	0.123	0.436***	-0.325***	-0.197***	1.000				
9. LEV	0.00	-0.047	-0.050	0.030	0.083	0.089	0.005	-0.068	1.000			
10. ROA	0.53***	0.161**	0.023	0.311***	0.231***	0.102	-0.003	0.374***	-0.145**	1.000		
11. ROE	0.28***	0.009	0.029	0.093	0.135**	0.006	0.036	0.265***	-0.224***	0.357***	1.000	
12. MTBV	0.07	0.017	-0.025	0.191***	0.055	0.045	0.187***	0.096	0.057	0.203***	-0.484***	1.000

Table 6 represents the descriptive statistics of the predictor variables across the dataset of 230 firm-year observations, whilst table 7 indicates the correlation between them.

Table 6 | Descriptive statistics for dependent, independent, control variables and moderators

Variable	Observations	Mean	Minimum	Maximum	S.D.
OPE	230	0.75	0.62	1.00	0.09
IE	230	0.62	0.50	1.00	0.16
PE	230	0.92	0.69	1.00	0.08
SE	230	0.82	0.62	1.00	0.06
ESG	230	63.90	21.49	93.26	14.35
KC	230	59.03	38.30	63.70	4.62
DTI	230	0.00	-0.77	6.58	1.00
LnFS	230	16.11	12.24	19.79	1.69
LEV	230	23.01	0.00	91.37	17.06
ROA	230	11.16	-37.50	61.50	11.86
ROE	230	17.08	-1271.67	162.54	92.21
MTBV	230	7.44	-153.79	190.54	18.21

As the hypothesis 1 proposes that the relationship between ESG performance and firm efficiencies is negative, our results in table 8 partially support H1. There are negative associations between ESG and OPE (β = -0.00267, $p < 0.01$), PE (β = -0.0275, $p < 0.05$) and SE (β = -0.00479, $p < 0.01$). Our results are in line with the

findings of Duque-Grisales and Aguilera-Caracuel (2021), Gao et al. (2025), and Fisher-Vanden and Thorburn (2011).

Regarding two-way interaction between ESG performance and moderators, the moderating effect of digital transformation on the relationship between ESG and all dimensions of efficiency (ESG x DTI) is insignificant. Hence, our findings do not provide empirical support for H2. .

Finally, with respect to the three-way interaction (ESG x DTI x KC), the analysis reveals significant negative interaction effects for OPE (β = -0.000233, p < 0.1) and SE (β = -0.000265, p < 0.05). Our previous two-way interaction result, combined with this finding, suggest that digital transformation effects can become apparent in the presence of knowledge, since the integration of digital technologies, such as big data and IoT, with multidisciplinary knowledge, offers strategic capability in business sectors (Huesig and Endres, 2019; Scuotto et al., 2017; Singh and El-Kassar, 2019). However, the result is negative, meaning that excessive country-level knowledge capability can weaken the moderating effect of digital transformation on the relationship between ESG performance and both operational and sustainable efficiencies; accordingly, our findings provide partial support for H3. (Guo et al., 2023; Feliciano-Cestero et al., 2023; Zhong and Ren, 2023).

Table 8 | Truncated regression analysis of the effect of ESG on the performance with moderators.

	OPE	IE	PE	SE
Control variables				
LnFS	0.0422***	0.0364***	0.0931***	0.0490***
	(0.000735)	(0.00112)	(0.0286)	(0.00105)
LEV	0.000659*	0.000156	0.00569	0.000774*
	(0.000382)	(0.000445)	(0.00564)	(0.000453)
ROA	0.00321***	0.00146*	0.0123	0.000853
	(0.000663)	(0.000885)	(0.0100)	(0.000628)
ROE	-0.0000756	-0.000131**	0.00177	0.0000532
	(0.0000791)	(0.0000645)	(0.00167)	(0.0000766)
MTBV	0.000170	-0.000742**	0.0139	0.000553
	(0.000579)	(0.000358)	(0.0131)	(0.000413)
Independent variables				
ESG	-0.00267***	-0.000207	-0.0275**	-0.00479***
	(0.000456)	(0.000714)	(0.0116)	(0.000479)
Moderators				
DTI	-0.0220***	0.00528	-0.125*	-0.0374***
	(0.00609)	(0.00982)	(0.0701)	(0.00583)
KC	0.00129	-0.00129	0.0122	-0.000161
	(0.00160)	(0.00260)	(0.0186)	(0.00191)
2-way interaction				
ESGxDTI	0.000271	-0.000648	-0.00320	-0.000274
	(0.000654)	(0.000752)	(0.00545)	(0.000546)
ESGxKC	0.000175**	-0.000210	-0.00191	0.000144
	(0.0000887)	(0.000134)	(0.00117)	(0.000131)
DTIxKC	-0.00293*	0.000290	0.00862	-0.00346**
	(0.00150)	(0.00268)	(0.0171)	(0.00155)
3-way interaction				
ESGxDTIxKC	-0.000233*	0.000396	-0.00221	-0.000265**
	(0.000134)	(0.000253)	(0.00195)	(0.000131)
Log likelihood	253.33	162.03	282.94	241.91
Wald squared	28044.68***	8597.74***	42.05***	26589.06***

Standard errors in parentheses
* *p*<0.1, ** *p*<0.05, *** *p*<0.01

5 Conclusion and Recommendation

5.1 Conclusions

This study examines how ESG performance relates to corporate efficiency in the technology sector and how this relationship is shaped by digital transformation and country-level knowledge capability. Using a sample of 46 NASDAQ-listed firms over the period 2019–2023.

ESG engagement is associated with lower short-term efficiency across most dimensions, suggesting that adjustment costs and resource reallocation dominate in the early stages.

DTI, on its own, does not alter this relationship in a meaningful way. Its effect becomes more apparent only when combined with higher levels of KC. In such settings, DTI and KC jointly strengthen the negative association between ESG and efficiency.

5.2 Limitation and Further Research

This study has several limitations that provide directions for future research. First, the study uses an aggregate ESG score, which may mask differences across environmental, social, and governance dimensions. Further research could explore how these dimensions individually relate to corporate efficiency and how digital transformation and knowledge capability influence these relationships.

Second, while the textual analysis method captures firms' strategic emphasis on digitalization, it may not fully reflect the depth or effectiveness of actual digital implementation. Future research could adopt more refined measures to capture the quality and outcomes of digital transformation.

References

Ahmad, N., Mobarek, A., & Roni, N. N. (2021). Revisiting the impact of ESG on financial performance of FTSE350 UK firms: Static and dynamic panel data analysis. Cogent Business & Management, 8(1), 1900500. https://doi.org/10.1080/23311975.2021.1900500

Brunetti, F., Matt, D. T., Bonfanti, A., De Longhi, A., Pedrini, G., & Orzes, G. (2020). Digital transformation challenges: Strategies emerging from a multi-stakeholder approach. The TQM Journal, 32(4), 697–724. https://doi.org/10.1108/TQM-12-2019-0309

Cao, Q., Zhu, T., & Yu, W. (2024). ESG investment and bank efficiency: Evidence from China. Energy Economics, 133, 107516.

Chang, K., Cheng, X., Wang, Y., Liu, Q., & Hu, J. (2023). The impacts of ESG performance and digital finance on corporate financing efficiency in China. Applied Economics Letters, 30(4), 516–523. https://doi.org/10.1080/13504851.2021.1996527

Chanias, S., Myers, M. D., & Hess, T. (2019). Digital transformation strategy making in pre-digital organizations: The case of a financial services provider. The Journal of Strategic Information Systems, 28(1), 17–33.

Cohen, W. M., & Levinthal, D. A. (1990). Absorptive capacity: A new perspective on learning and innovation. Administrative Science Quarterly, 35(1), 128–152.

Crupi, A., Del Sarto, N., Di Minin, A., Gregori, G. L., Lepore, D., Marinelli, L., & Spigarelli, F. (2020). The digital transformation of SMEs – a new knowledge broker called the digital innovation hub. Journal of Knowledge Management, 24(6), 1263–1288. https://doi.org/10.1108/JKM-11-2019-0623

Duque-Grisales, E., & Aguilera-Caracuel, J. (2021). Environmental, Social and Governance (ESG) Scores and Financial Performance of Multilatinas: Moderating Effects of Geographic International Diversification and Financial Slack. Journal of Business Ethics, 168(2), 315–334. https://doi.org/10.1007/s10551-019-04177-w

Endrikat, J., De Villiers, C., Guenther, T. W., & Guenther, E. M. (2021). Board Characteristics and Corporate Social Responsibility: A Meta-Analytic Investigation. Business & Society, 60(8), 2099–2135. https://doi.org/10.1177/0007650320930638

Feliciano-Cestero, M. M., Ameen, N., Kotabe, M., Paul, J., & Signoret, M. (2023). Is digital transformation threatened? A systematic literature review of the factors influencing firms' digital transformation and internationalization. Journal of Business Research, 157, 113546.

Fernández Sánchez, J. L., Luna Sotorrío, L., & Baraibar Diez, E. (2015). The relationship between corporate social responsibility and corporate reputation in a turbulent environment: Spanish evidence of the Ibex35 firms. Corporate Governance, 15(4), 563–575.

Fisher-Vanden, K., & Thorburn, K. S. (2011). Voluntary corporate environmental initiatives and shareholder wealth. Journal of Environmental Economics and management, 62(3), 430-445.

Freeman, R. E. (2010). Strategic management: A stakeholder approach. Cambridge university press. https://books.google.com/books?hl=en&lr=&id=NpmA_qEiOpkC&oi=fnd&pg=PR5&dq=Strategic+management:+A+stakeholder+approach&ots=62hIL8P5PO&sig=23Wbve-bbc8VZNyvj38bzpZo1tc

Friedman, M. (2007). The social responsibility of business is to increase its profits. In Corporate ethics and corporate governance (pp. 173–178). Springer. https://link.springer.com/chapter/10.1007/978-3-540-70818-6_14

Gao, Y., Omenihu, C. M., Brahma, S., & Nwafor, C. (2025). Short-term impact of ESG performance on default risk under the green transition of energy sector: Evidence in China. Administrative Sciences, 15(9), 352.

Ge, T., Li, J., & Wang, C. (2023). Econometric analysis of the impact of innovative city pilots on CO2 emissions in China. Environment, Development and Sustainability, 25(9), 9359–9386. https://doi.org/10.1007/s10668-022-02439-8

Gebauer, H., Fleisch, E., Lamprecht, C., & Wortmann, F. (2020). Growth paths for overcoming the digitalization paradox. Business Horizons, 63(3), 313–323.

Guo, X., Li, M., Wang, Y., & Mardani, A. (2023). Does digital transformation improve the firm's performance? From the perspective of digitalization paradox and managerial myopia. Journal of Business Research, 163, 113868.

Han, X., Hui, N., & Song, W. (2014). Can informatization improve the technology innovation efficiency of Chinese industrial sectors. China Ind. Econ, 12, 70-82.

Hassel, L., Nilsson, H., & Nyquist, S. (2005). The value relevance of environmental performance. European Accounting Review, 14(1), 41-61.

Horváth, D., & Szabó, R. Z. (2019). Driving forces and barriers of Industry 4.0: Do multinational and small and medium-sized companies have equal opportunities? Technological Forecasting and Social Change, 146, 119–132.

Huesig, S., & Endres, H. (2019). Exploring the digital innovation process: The role of functionality for the adoption of innovation management software by innovation managers. European Journal of Innovation Management, 22(2), 302–314.

Kandpal, V., Jaswal, A., Santibanez Gonzalez, E. D. R., & Agarwal, N. (2024). Corporate Social Responsibility (C.S.R.) and E.S.G. Reporting: Redefining Business in the Twenty-First Century. In V. Kandpal, A. Jaswal, E. D. R. Santibanez Gonzalez, & N. Agarwal, Sustainable Energy Transition (pp. 239–272). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-52943-6_8

Kriebel, J., & Debener, J. (2019). The effect of digital transformation on bank performance. Available at SSRN 3461594. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3461594

Le, M. H., & Lu, W. M. (2025). The Moderating Effect of Multinationality on Environmental, Social, and Governance and Efficiency: A Strategic Analysis of Multinational Enterprises. Business Strategy and the Environment, 34(8), 10723–10740. https://doi.org/10.1002/bse.70126

Li, X., Xu, K., Wang, Q., & Wang, Y. (2025). Digital transformation and corporate innovation: the trade-off between financing obtained and risk increasing. Journal of Intelligent Manufacturing and Special Equipment, 6(3), 225-247.

London Stock Exchange (LSEG). 2025. LSEG Data & Analytics. LSEG.

Malik, M. (2015). Value-Enhancing Capabilities of CSR: A Brief Review of Contemporary Literature. Journal of Business Ethics, 127(2), 419–438. https://doi.org/10.1007/s10551-014-2051-9

Matarazzo, M., Penco, L., Profumo, G., & Quaglia, R. (2021). Digital transformation and customer value creation in Made in Italy SMEs: A dynamic capabilities perspective. Journal of Business Research, 123, 642–656.

Patten, D. M. (2002). The relation between environmental performance and environmental disclosure: a research note. Accounting, organizations and Society, 27(8), 763-773.

Pham, T. N., Tran, P. P., Le, M.-H., Vo, H. N., Pham, C. D., & Nguyen, H.-D. (2022). The effects of ESG combined score on business performance of enterprises in the transportation industry. Sustainability, 14(14), 8354.

Preston, L. E., & O'Bannon, D. P. (1997). The Corporate Social-Financial Performance Relationship: A Typology and Analysis. Business & Society, 36(4), 419–429. https://doi.org/10.1177/000765039703600406

Qi, Y., Chen, Q., Yang, M., & Sun, Y. (2024). Ambidextrous knowledge accumulation, dynamic capability and manufacturing digital transformation in China. Journal of Knowledge Management, 28(8), 2275–2305.

Roberts, N., Galluch, P. S., Dinger, M., & Grover, V. (2012). Absorptive capacity and information systems research: Review, synthesis, and directions for future research. MIS Quarterly, 625–648.

Scuotto, V., Santoro, G., Bresciani, S., & Del Giudice, M. (2017). Shifting intra- and inter-organizational innovation processes towards digital business: An empirical analysis of SMEs. Creativity and Innovation Management, 26(3), 247–255. https://doi.org/10.1111/caim.12221

Shakina, E., Parshakov, P., & Alsufiev, A. (2021). Rethinking the corporate digital divide: The complementarity of technologies and the demand for digital skills. Technological Forecasting and Social Change, 162, 120405.

Singh, S. K., & El-Kassar, A.-N. (2019). Role of big data analytics in developing sustainable capabilities. Journal of Cleaner Production, 213, 1264–1273.

Wang, L., & Yang, H. (2024). Digital technology innovation and corporate ESG performance: Evidence from China. Economic Change and Restructuring, 57(6), 207. https://doi.org/10.1007/s10644-024-09791-x

Wanke, P., Ostovan, S., Mozaffari, M. R., Gerami, J., & Tan, Y. (2023). Stochastic network DEA-R models for two-stage systems. Journal of Modelling in Management, 18(3), 842–875. https://doi.org/10.1108/JM2-10-2021-0256

World Intellectual Property Organisation (WIPO). 2023. Global Innovation Index 2023: Innovation in the Face of Uncertainty. WIPO. https:// doi. org/ 10. 34667/ tind. 48220

Xie, J., Nozawa, W., Yagi, M., Fujii, H., & Managi, S. (2019). Do environmental, social, and governance activities improve corporate financial performance? Business Strategy and the Environment, 28(2), 286–300. https://doi.org/10.1002/bse.2224

Xiong, X. (2024). Examining the influence of knowledge transfer and dynamic capabilities on enterprise digital transformation. PloS One, 19(12), e0311176.

Xu, P., Chen, L., & Dai, H. (2022). Pathways to sustainable development: Corporate digital transformation and environmental performance in China. Sustainability, 15(1), 256.

Yu, A., Shi, Y., You, J., & Zhu, J. (2021). Innovation performance evaluation for high-tech companies using a dynamic network data envelopment analysis approach. European Journal of Operational Research, 292(1), 199–212.

Zhong, X., & Ren, G. (2023). Independent and joint effects of CSR and CSI on the effectiveness of digital transformation for transition economy firms. *Journal of Business Research*, *156*, 113478.

Developing an Intelligent Organizational Carbon Inventory Assistant Integrating LangChain and Large Language Models: An Empirical Study

Huang, Yan-Ping
Graduate Institute of Information and Decision Sciences, National Taipei University of Business, Taipei, Taiwan

Abstract

Continued advances in manufacturing technology have improved production precision but have also intensified resource consumption and carbon emissions, making corporate greenhouse gas (GHG) management an urgent priority. Carbon inventory is a foundational tool for this purpose, enabling firms to evaluate their carbon footprint and formulate emission-reduction strategies aligned with climate goals. In practice, however, organizations face substantial barriers when implementing carbon inventories, including difficulties in interpreting accounting standards, complex data collection, and heavy reliance on specialized expertise and external consultants. This study investigates the practices, experiences, and challenges of organizational carbon inventory and management, and examines how generative artificial intelligence can lower these knowledge barriers. Using company A—a manufacturer of medical, 3C, and automotive modification components in Taoyuan, Taiwan—as a case study, an organizational-level GHG inventory was conducted in accordance with the ISO 14064-1 standard to identify emission sources and quantify the firm's 2022 carbon footprint. Building on this inventory, the study developed a carbon-knowledge intelligent assistant by integrating the LangChain framework with a large language model (GPT-3.5), applying a retrieval-augmented generation mechanism grounded in local carbon-inventory documents, emission-factor databases, and standard operating procedures. To evaluate the system, the LangChain-integrated model was compared with standalone ChatGPT-3.5 across five representative carbon-inventory question-answering tasks. A total of 39 participants with practical experience in or active involvement with carbon inventory implementation were invited to assess both approaches in terms of knowledge accuracy and ease of understanding. The results show that the LangChain-based system consistently outperformed standalone ChatGPT-3.5, achieving an average correctness rate of 61.2% compared with 38.8%—an overall improvement of 22.4 percentage points—while also demonstrating an advantage in ease of understanding. These findings confirm that grounding a large language model in domain-specific knowledge sources markedly improves the professionalism, consistency, and reliability of generated responses in specialized applications. Overall, the study provides both empirical evidence and practical insights for organizations seeking to adopt generative artificial intelligence to strengthen their carbon inventory and carbon management practices, and it contributes to the emerging literature on applying large language models as knowledge-support tools in sustainability management.

Keywords: Autonomous AI Agent, Large Language Model, Chatbots, Intelligent Assistant, Organizational Carbon Accounting, Carbon Knowledge.

1. Research Motivation and Objectives

1.1 Research Motivation

Climate change has become one of the most pressing challenges to global sustainable development. Rising average temperatures and increasingly frequent extreme weather events threaten human societies, industry, and ecosystems. In response, the international community has established climate governance frameworks—most notably the United Nations Framework Convention on Climate Change (UNFCCC) and the Kyoto Protocol—that lay the institutional foundation for greenhouse gas mitigation and extend decarbonization responsibilities to both governments and firms.

Within this context, carbon market mechanisms are widely recognized as an effective instrument for improving the economic efficiency of emissions reduction. By enabling the trading of emission allowances, carbon markets allow firms to adopt cost-effective mitigation strategies while collectively meeting reduction targets. Their effective functioning, however, depends on accurate, consistent, and transparent emissions data, making corporate carbon inventories a critical prerequisite for both carbon management and market participation.

A corporate carbon inventory is the systematic quantification of an organization's greenhouse gas emissions over a defined period according to recognized standards. It provides essential insight into emission structures and forms the basis for reduction planning. In practice, many firms encounter significant challenges—difficulties

interpreting accounting standards, complex data collection and integration, and high demands for specialized expertise. These barriers foster heavy reliance on external consultants and limit scalability, particularly among small and medium-sized enterprises.

Recent advances in generative artificial intelligence and large language models (LLMs) offer new opportunities to support such knowledge-intensive tasks. While existing studies have mainly examined the general question-answering capabilities of LLMs, empirical research on intelligent assistants designed specifically for organizational carbon inventory and management knowledge remains limited. This study is therefore motivated by the need to integrate carbon inventory practice with generative AI in order to lower knowledge barriers, improve the accessibility and interpretability of carbon-related information, and help enterprises prepare for carbon-market participation and long-term sustainability.

1.2 Research Objectives

The primary objective of this study is to apply organizational carbon inventory methodologies to systematically analyze the carbon footprint of corporate activities and to examine their value for carbon management and sustainable-development decision-making. The study also addresses a gap in the literature by developing and empirically evaluating a carbon-knowledge question-answering system, with emphasis on how far LLMs can facilitate enterprises' understanding of carbon inventory concepts and practices.

Company A, located in Taoyuan City, serves as the case study. An organizational-level carbon inventory is conducted according to relevant standards to analyze the sources and structure of emissions and to explore the role of carbon inventory in corporate carbon management and sustainability strategy. A carbon-knowledge intelligent assistant is then developed by integrating the LangChain framework with LLMs to provide decision-support and knowledge-consultation functions. The specific aims are summarized as three phases:

Phase One: To compile and systematize key carbon-inventory terminology and greenhouse gas accounting requirements, supporting on-site investigation and preparatory activities. Phase Two: To implement a carbon-knowledge assistant integrating LangChain with a ChatGPT-based model for natural language processing, automated question answering, and text generation, and to assess its value for enterprises planning inventories. Phase Three: To conduct an experimental evaluation comparing the correctness of carbon-knowledge responses between the GPT-3.5 model integrated with LangChain and the conventional ChatGPT-3.5 model.

2. Literature Review

2.1 Organizational Carbon Inventory and Carbon Management

As global climate governance frameworks have become institutionalized, organizational-level carbon inventories have emerged as a critical foundation for corporate carbon management and sustainability decision-making. Prior studies show that systematic inventory practices allow firms to identify major emission sources and to use inventory results as a basis for reduction strategy and performance evaluation (Schaltegger & Csutora, 2012). Carbon inventories thus function not merely as compliance tools but as essential informational infrastructure for internal management and stakeholder communication.

At the international level, ISO 14064-1 is widely adopted as the core framework for organizational GHG accounting, offering explicit guidance on boundary setting, emission-source classification, and data quality to enhance the consistency, transparency, and comparability of results (ISO, 2018). Standardized procedures also strengthen the credibility of carbon information in sustainability reporting, disclosure initiatives, and carbon-market participation. Nevertheless, the literature consistently reports substantial implementation challenges, including data-collection difficulties, cross-departmental coordination costs, and the expertise required to apply accounting standards (Burritt et al., 2011; Ascui & Lovell, 2012). Small and medium-sized enterprises, lacking resources and specialized personnel, are especially constrained—highlighting the pressing need to reduce the knowledge and operational barriers associated with carbon inventory.

2.2 Generative AI and Knowledge-Oriented Question-Answering Systems

Recent advances in generative artificial intelligence (GAI) and large language models (LLMs) provide new pathways for supporting knowledge-intensive tasks. Empirical research shows that LLMs achieve near-human performance in natural language understanding, text generation, and question answering (Brown et al., 2020; OpenAI, 2023; Huang et al., 2026), and they are increasingly deployed in customer service, knowledge management, and decision support to improve information accessibility (Dwivedi et al., 2023). However, the reliability and interpretability of LLMs in domain-specific contexts remain contested (Bender et al., 2021).

Because language models generate responses probabilistically, they may produce inaccurate or inconsistent outputs when not grounded in authoritative knowledge (Ji et al., 2023). Recent work therefore emphasizes integrating LLMs with external knowledge bases to enhance the trustworthiness and stability of their outputs in professional applications.

2.3 LangChain Architecture

LangChain is an application framework designed for large language model deployment, characterized by a modular architecture that integrates language models, external data sources, and reasoning workflows to enable context-aware retrieval and task-oriented applications (Chase, 2022). It provides tools, components, and interfaces that simplify the development of LLM-supported applications such as AI chatbots, generative question-answering systems, and summarization tools. Prior research on retrieval-augmented generation (RAG) has demonstrated its potential in law, healthcare, and enterprise knowledge management, where answer consistency and knowledge traceability are critical (Lewis et al., 2020; Karpukhin et al., 2020).

As illustrated in Figure 1, the LangChain architecture forms an end-to-end pipeline that transforms local unstructured documents into context-aware LLM responses (Chase, Maass, & Gan, 2023). Document loaders extract text from local sources; text splitters divide the content into manageable chunks; each chunk is converted into vector embeddings and stored in a vector database. When a user submits a query, it is likewise embedded and compared against stored embeddings using similarity search. Relevant chunks are retrieved and combined with predefined prompt templates to construct a structured input, and the LLM then generates a response grounded in both the retrieved context and the query. This RAG process improves the relevance, accuracy, and coherence of outputs by explicitly integrating external knowledge into the inference workflow.

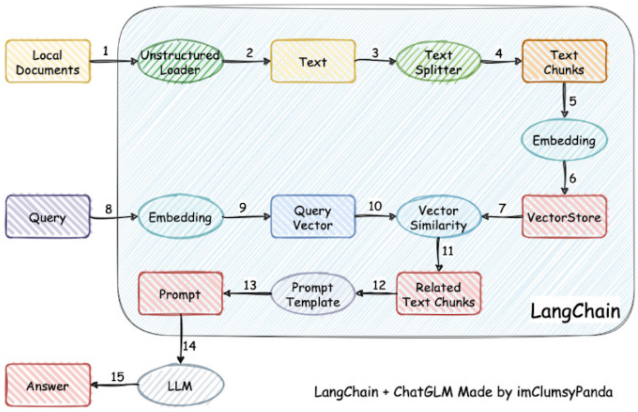


Figure 1: Algorithm architecture of LangChain (Chase et al., 2023).

Despite this potential, empirical research on intelligent carbon-knowledge assistants that integrate LangChain with LLMs remains scarce, and comparative evaluations of model architectures with respect to knowledge accuracy and perceived ease of understanding are largely absent. This study addresses that gap through an empirical design evaluating the practical effectiveness of generative-AI knowledge-support systems in organizational carbon inventory contexts.

3. Research Methodology

This chapter describes the research design and implementation in three parts: the research subject and scope, the research instruments, and the research framework and procedures.

3.1 Research Subject and Scope

The research subject is Company A, established in 2021 and engaged in developing and manufacturing medical products, 3C electronics, automotive modification components, and related parts. Its core technologies include closed-die forging of magnesium, aluminum, and lithium alloys. Responding to rising demand for lightweight materials in new-energy vehicles and unmanned transportation, the company invests in R&D to provide lightweight, high-strength, energy-efficient solutions, and it holds ISO 9001:2015 certification. Its principal markets—Europe, North America, and the United Kingdom—impose stringent environmental and sustainability requirements. Accordingly, the company has adopted measures such as rooftop ventilation to

improve heat dissipation and energy efficiency, and low-waste manufacturing under which roughly 95% of raw materials are effectively utilized and the remaining 5% recycled.

In response to the EU Carbon Border Adjustment Mechanism (CBAM), which requires carbon-content declaration for imported products from October 2023, and to proposed U.S. carbon-tariff legislation under the Clean Competition Act, Company A has begun planning carbon-inventory systems. With governmental support, it is conducting organizational-level GHG inventories and personnel training to understand its current emission profile, formulate feasible reduction pathways, meet customer and regulatory demands, and enhance competitiveness. To this end, the company convened all relevant departments to form a carbon-inventory task force, and this study provided the participating personnel with essential GHG education and training covering kick-off, organizational-boundary and emission-source identification, quantification and calculation, documentation and reporting, and internal verification.

3.2 Research Framework

After data collection and organization, this study integrates the LangChain framework with large language models to construct a question-and-answer chatbot for organizational carbon knowledge. The framework combines GHG inventory outcomes with the semantic understanding and generative capabilities of LLMs to enhance the accessibility and practical value of inventory information. The overall workflow comprises: (1) defining inventory objectives and scope; (2) identifying inventory boundaries and emission sources; (3) collecting emission-related data from internal records, government data, and third-party datasets; (4) processing data through cleaning, organization, and calibration; (5) calculating emissions using appropriate methods and emission factors; (6) importing results and documentation into LangChain, where text segmentation generates semantically meaningful chunks and keywords; and (7) using the LLM to generate a carbon-knowledge Q&A repository and an interactive chatbot.

3.2.1 Carbon Inventory Method

This study adopts the organizational carbon inventory methodology of ISO 14064-1, an international standard supporting the accurate, transparent, and consistent quantification, reporting, and verification of GHG emissions and removals. Its procedure follows a four-step process—boundary setting, source identification, calculation, and reporting—summarized in Table1.

Table 1: Four Steps of the Carbon Inventory Process

GHG Inventory Step	Description
Boundary Setting	Defines the foundation of the inventory based on organizational scale and type, activity nature, and emission profile. Boundaries are classified into organizational boundaries (the entities covered) and emission-source boundaries (the sources included).
Emission Source Identification	Identifies all direct and indirect emission sources within the boundary. Direct sources (e.g., stationary combustion, process, mobile, and fugitive emissions) are under organizational control; indirect sources (e.g., purchased energy, waste, transportation, commuting) are not.
Emission Calculation	The core step: total emissions are quantified using established GHG calculation methodologies.
Reporting	The primary output: a carbon inventory report disclosing GHG emission information to external stakeholders.

Source: Climate Change Administration, Ministry of Environment, Taiwan (2024).

3.2.2 Greenhouse Gas Emission Standards

Because many domestic enterprises are integral to international supply chains, Taiwan's Ministry of Environment revised its Greenhouse Gas Inventory Guidelines, officially released in February 2024, to provide a cross-industry reference for inventory implementation. The Guidelines correspond to Categories 1–6 of ISO 14064-1:2018, with adjustments for domestic regulatory practice. Table 2 summarizes the main correspondences between the international standard and the national Guidelines, which serve as the basis for subsequent system development and data integration.

Table 2: Comparison of Greenhouse Gas Emission Standards

Item	ISO 14064-1:2018	National GHG Inventory Guidelines
Issuing Organization	International Organization for Standardization (ISO)	Ministry of Environment, Taiwan
Scope of Application	International verification and voluntary GHG inventories	Entities regulated by the Ministry of Environment and the Financial Supervisory Commission
Classification	Scopes 1–6 (direct; energy indirect; transportation; products used; products from the organization; other indirect)	Scopes I–III, subdivided into stationary, process, mobile, fugitive, land-use, purchased energy, transportation, commuting, and product-related sources
Measurement Unit	Metric tons of CO_2 equivalent (tCO_2e)	Metric tons of CO_2 equivalent (tCO_2e)

Source: Climate Change Administration, Ministry of Environment, Taiwan (2024).

3.3 LangChain and Large Language Model Integration

This study adopts the LangChain framework (Chase et al., 2023) with OpenAI's GPT-3.5 model to construct an intelligent assistant. LangChain provides strong text-comprehension and semantic-extraction capabilities for retrieving carbon-related knowledge, while GPT-3.5 offers generative capabilities for automatically producing carbon-related question-and-answer pairs (Figure 2). The assistant is developed by (1) collecting and organizing carbon-related knowledge from government reports, journal articles, and online resources into a local PDF knowledge base; (2) using LangChain to interpret semantics and extract key information; and (3) using GPT-3.5 to generate carbon-related Q&A pairs.

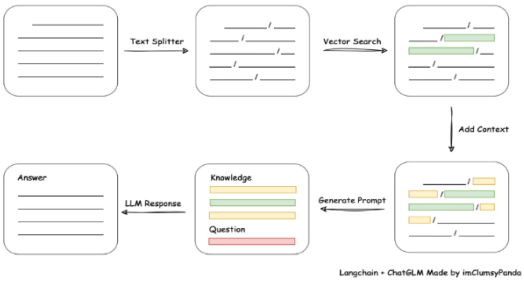


Figure 2: Workflow of LangChain and large language model integration (Chase et al., 2023).

For example, from the statement “Carbon emissions refer to greenhouse gas emissions generated by human activities,” LangChain extracts key elements (carbon emissions are GHG emissions; the source is human activities; greenhouse gases are a category of gases) and generates questions such as “What are carbon emissions?” and “How are carbon emissions generated?” GPT-3.5 then produces corresponding answers, and the resulting Q&A pairs are used to evaluate the assistant's performance.

3.4 Research Design

The evaluation design draws on the conceptual framework of the Turing Test, proposed by Alan Turing in his 1950 paper “Computing Machinery and Intelligence,” in which a human evaluator attempts to distinguish machine-generated from human responses through text-based interaction. Adapting this logic, the study adopts a comparative evaluation design in which participants interested in or engaged with carbon inventory assess responses generated by two systems—LangChain integrated with GPT-3.5, and standalone ChatGPT-3.5—both of which exhibit advanced natural-language capabilities. This Turing-Test-inspired approach examines the extent to which LLM-based carbon-knowledge assistants can produce responses perceived as accurate, natural, and professionally reliable in carbon inventory contexts.

4. Research Results

4.1 Organizational Carbon Inventory Assessment and Results

Following ISO 14064-1:2018 and the GHG Protocol (WBCSD and WRI), the study applies an operational-control approach to set the inventory boundary and consolidates Company A's GHG emissions and removals for 2022 (Figure 3). The inventory covers the company's operational boundaries and systematically identifies direct

emissions, energy-related indirect emissions, and other indirect emissions, encompassing carbon oxides and fluorinated gases per the classification of the Climate Change Administration, Ministry of Environment, Taiwan (2024).

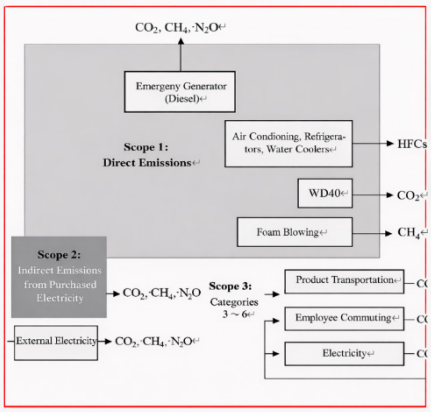


Figure3: Boundary-setting methodology for the greenhouse gas inventory.

4.1.1 Direct GHG Emission Sources (Scope 1)

Direct emissions originate from sources under the organization's operational control. For Company A, identified sources include emergency generators, lubricants (e.g., WD-40), refrigerant leakage from refrigeration and air-conditioning equipment, water dispensers, restroom septic tanks, and fire extinguishers. During the inventory year the emergency generator was neither procured nor operated, and certain equipment (the Weston cooling unit, water dispensers, R-22 and R-600a refrigerants, and ABC dry-powder fire extinguishers) was identified but excluded from quantification owing to missing charge information or applicable guidelines. Results are summarized in Table 3.

Table 3: Direct Greenhouse Gas Emissions (Scope 1)

Category	Emissions (tCO_2e)	Percentage (%)
Category 1: Direct GHG emissions and removals	2.8867	100%
1.1 Stationary combustion	0	0%
1.2 Mobile combustion	0	0%
1.3 Process emissions	0	0%
1.4 Fugitive emissions	2.8867	100%

4.1.2 Energy-Related Indirect GHG Emission Sources (Scope 2)

Company A's Scope 2 emissions derive primarily from purchased electricity supplied by Taiwan Power Company. Based on the 2022 grid emission factor of 0.495 kg CO_2e per kWh announced by the Bureau of Energy, Ministry of Economic Affairs, total Scope 2 emissions were calculated as shown in Table 4.

Table 4: Energy-Related Greenhouse Gas Emissions (Scope 2)

Category	Emissions (tCO_2e)	Percentage (%)
Category 2: Energy-related indirect GHG emissions	86.6448	100%
2.1 Purchased electricity	86.6448	100%

4.1.3 Other Indirect GHG Emission Sources (Scope 3)

Under the GHG Protocol, Scope 3 covers indirect emissions from activities such as employee commuting, upstream transportation and distribution of raw materials, and fuel- and energy-related activities. Following ISO 14064-1:2018, these are reported under Categories 3 and 4, which together amounted to 134.4287 tCO_2e (Tables 5 and 6). Company A also operates six landlord-provided window air conditioners whose R-22 refrigerant leakage falls under Category 6; these were identified but excluded from the material indirect categories per the inventory guidelines.

Table 5: Category 3 Greenhouse Gas Emissions

Category	Emissions (tCO_2e)	Percentage (%)
Category 3: Indirect GHG emissions – Transportation	118.9902	100%
3.1 Upstream transportation and distribution	108.3929	91.09%
3.3 Employee commuting	10.5973	8.91%

Table 6: Category 4 Greenhouse Gas Emissions

Category	Emissions (tCO_2e)	Percentage (%)
Category 4: Indirect GHG emissions – Use of products by the organization	15.4385	100%
4.1 Upstream purchased electricity	15.4385	100%

4.1.4 Emission Calculation and Quantification

The emission-factor approach was used to quantify emissions. Activity data were collected using the Greenhouse Gas Inventory Tool, and emission factors were obtained from the Greenhouse Gas Emission Factor Management Table (Version 6.0.4) (Ministry of Environment, 2019). Because gases differ in climate impact, emissions were converted into CO_2 equivalents using Global Warming Potential (GWP) values from the IPCC Sixth Assessment Report (AR6, 2021). The principal calculation formulas are: stationary and mobile combustion, GHG emissions = fuel consumption × emission factor × GWP; fugitive emissions from septic tanks, emissions = total working hours × septic-tank emission factor × CH_4 GWP; refrigerant leakage, emissions = annual leakage amount × GWP; purchased electricity, CO_2e = electricity consumption × emission factor (0.495 kg CO_2e/kWh); upstream transportation, emissions = vehicle emission factor ($kgCO_2e$/tkm) × cargo weight (t) × distance (km); employee commuting, emissions = transport emission factor ($kgCO_2e$/pkm) × one-way distance (km) × 2 × working days; and upstream electricity-related emissions, using the indirect factor of 0.0882 kg CO_2e/kWh (2020). Total emissions are summarized in Table 7.

Table 7: Total Greenhouse Gas Emissions of Company A in 2022

Scope	Category	Emissions tCO_2e)	Share (%)	Notes
Scope 1	Fugitive emissions	2.8867	1.29%	Stationary/mobile/process = 0
Scope 2	Purchased electricity	86.6448	38.69%	Taipower grid factor 0.495
Scope 3	Upstream transportation & distribution	108.3929	48.40%	Largest single source
Scope 3	Employee commuting	10.5973	4.73%	
Scope 3	Upstream electricity-related indirect	15.4385	6.89%	
Total		223.9602	100.00%	

4.2 Results of the Q&A Chatbot Using LangChain

An organizational carbon-inventory intelligent assistant was developed using OpenAI LLMs, Python, and LangChain to integrate domain-knowledge retrieval and question answering. Carbon-related knowledge was stored locally, and the resulting chatbot was applied within the case company to provide guidance on practical challenges (Figures 4 to 6), including: how fire extinguishers should be inventoried and which fall within the boundary; how to address data limitations for methane-generating septic tanks when design and operational records are incomplete; how to define and calculate emissions from employee business travel; and how to select appropriate emission factors, whose choice directly affects the accuracy of results. Overall, the LangChain-based chatbot provided structured explanations and practical guidance, enabling the company to conduct its inventory more systematically and confidently.

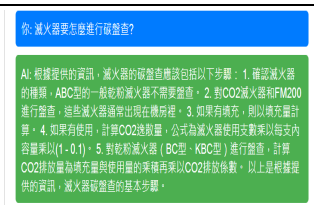


Figure 4: Fire-extinguisher inventory within the boundary.

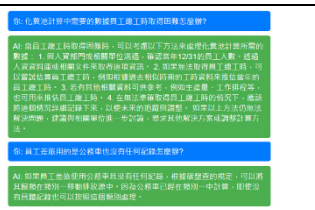


Figure 5: Practical challenges in the inventory process.

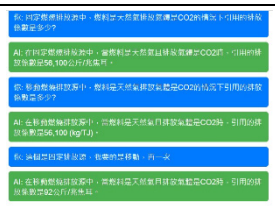


Figure 6: Emission factors used in quantification.

4.3 Questionnaire Analysis of LangChain versus ChatGPT

The experiment recruited 39 participants interested in carbon inventory or carbon-related knowledge to compare the text-generation performance of LangChain integrated with GPT-3.5 against standalone ChatGPT-3.5 across five carbon-inventory tasks; 39 valid questionnaires were collected. Female participants accounted for 69.23% (n = 27) and males 30.77% (n = 12). The largest age group was 21–30 (38.5%, n = 15), followed by 41–50 and 51+ (25.6% each). In education, 97.4% held a bachelor's degree or higher, with master's and doctoral

holders together comprising 46.1%, indicating a well-qualified sample suitable for evaluating generative-AI applications in carbon inventory.

As shown in Table 8, LangChain integrated with GPT-3.5 consistently outperformed standalone ChatGPT-3.5 across all five dimensions, with the largest gap in the general-waste inventory method (70% vs. 30%). The average correctness rate of the LangChain system was (60+61+70+58+57)/5 = 61%, compared with (40+39+30+42+43)/5 = 39% for ChatGPT-3.5—an overall improvement of 22.4 percentage points.

Table 8: Comparison of Text-Generation Correctness between LangChain and ChatGPT

System	Fire extinguisher	Septic tank	General waste	Official vehicles (no records)	Natural gas factor
LangChain + GPT-3.5	60%	61%	70%	58%	57%
ChatGPT-3.5	40%	39%	30%	42%	43%

These results indicate that integrating a knowledge-retrieval mechanism improves response accuracy by roughly 22% in professional carbon-inventory question answering. By grounding responses in local inventory documents, emission-factor databases, and standard operating procedures, retrieval-augmented generation improves consistency and reliability, making such systems well suited to domain-specific applications like carbon inventory management.

5. Conclusions

5.1 Research Summary

This study confirms that a large language model-based carbon inventory assistant is an effective means of enhancing organizational understanding of carbon-inventory knowledge. By integrating LangChain with GPT-3.5, it designed and implemented a carbon-knowledge Q&A assistant that helps organizations rapidly grasp key concepts, including definitions, calculation methods, application contexts, and managerial value. The comparative evaluation with 39 experienced participants showed that, when sufficient domain knowledge is provided, the generated responses are perceived as more acceptable and easier to understand, demonstrating the practical potential of LLMs for supporting carbon-inventory knowledge and organizational learning.

5.2 Managerial Implications

The findings offer several managerial implications. First, an LLM-based assistant reduces the learning barriers of carbon inventory by providing immediate responses on definitions, calculation methods, and applicable scopes, lessening reliance on external consultants. Second, because inventory activities span production, equipment, procurement, logistics, and administration, the assistant serves as a shared knowledge tool that improves information consistency and supports more standardized processes. Third, it functions as both a decision-support tool—giving managers timely explanations of emission sources, scopes, and regulatory guidelines—and a training resource enabling non-specialist staff to acquire fundamental knowledge. Given the time-sensitivity of inventory regulations and emission factors, management should also establish regular data updates and response validation to ensure accuracy.

5.3 Future Research Directions

Several directions emerge. At the model level, future studies may adopt more advanced LLMs to compare knowledge comprehension, terminology handling, and response quality. Regarding knowledge sources, integrating additional structured and unstructured data—international emission databases, industry-specific factors, and regulatory documents—with automated update mechanisms would improve accuracy as standards are revised. At the functionality level, the assistant could extend beyond basic Q&A to incorporate emission-calculation modules, scenario simulation, and reduction-strategy recommendations. Finally, more comprehensive empirical evaluations using quantitative indicators and longitudinal studies across organizations of different sizes and industries would clarify the long-term impact of LLM adoption on carbon-management maturity.

6 Acknowledgements

The author thanks Engineer Lin Feng-Yu for technical support that made the successful completion of this study possible.

7 Ethics declaration

This research involved an anonymous questionnaire in which 39 adult participants voluntarily evaluated the accuracy and ease of understanding of system-generated responses. No personally identifiable, sensitive, or health-related data were collected, and participation was voluntary and anonymous. Under the applicable institutional guidelines, this type of low-risk, non-interventional survey did not require formal ethics committee clearance.

8 AI declaration

Generative AI tools form the subject of this study; the LangChain-integrated GPT-3.5 system and ChatGPT-3.5 were used as research instruments, and their outputs were evaluated by human participants as described in the methodology. All research design, data collection, analysis, interpretation, and conclusions are the author's own, and the author reviewed and verified all content to ensure its accuracy and integrity.

References

Ascui, F., & Lovell, H. (2012). Carbon accounting and the construction of competence. Journal of Cleaner Production, 36, 48–59.

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? Proceedings of the ACM Conference on Fairness, Accountability, and Transparency, 610–623.

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., … Amodei, D. (2020). Language models are few-shot learners. Advances in Neural Information Processing Systems, 33, 1877–1901.

Burritt, R. L., Schaltegger, S., & Zvezdov, D. (2011). Carbon management accounting: Explaining practice in leading German companies. Journal of Cleaner Production, 19(9–10), 829–836.

Chase, H. (2022). LangChain [Computer software]. GitHub.

Chase, H., Maass, A., & Gan, A. (2023). LangChain: Architecture and end-to-end pipelines for context-aware LLM applications. LangChain Documentation.

Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., … Raman, R. (2023). So what if ChatGPT wrote it? Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. MIS Quarterly, 47(1), 1–37.

Huang, Z., Long, R., Chen, H., Wu, M., Li, Q., & Na, J. (2026). LLMQR: Efficient knowledge distillation for questionnaire recommendation via large language models. Expert Systems with Applications, 303, 130632.

International Organization for Standardization. (2018). ISO 14064-1:2018 Greenhouse gases—Part 1: Specification with guidance at the organization level for quantification and reporting of greenhouse gas emissions and removals. ISO.

Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., … Fung, P. (2023). Survey of hallucination in natural language generation. ACM Computing Surveys, 55(12), Article 248.

Karpukhin, V., Oguz, B., Min, S., Lewis, P., Wu, L., Edunov, S., Chen, D., & Yih, W.-T. (2020). Dense passage retrieval for open-domain question answering. Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP), 6769–6781.

Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., … Kiela, D. (2020). Retrieval-augmented generation for knowledge-intensive NLP tasks. Advances in Neural Information Processing Systems, 33, 9459–9474.

Ministry of Environment, Climate Change Administration. (2019). Greenhouse gas emission factor management table (Version 6.0.4). Ministry of Environment, Taiwan.

Ministry of Environment, Climate Change Administration. (2024). Guidelines for greenhouse gas emissions inventory. Ministry of Environment, Taiwan.

OpenAI. (2023). ChatGPT (Large language model). https://chat.openai.com/

Schaltegger, S., & Csutora, M. (2012). Carbon accounting for sustainability and management. Journal of Cleaner Production, 36, 1–16.

Developing a Two-Dimensional Portfolio Matrix of Innovation and Profitability: Evidence from Solar Energy Firms

Huang, Yan-Ping[1] Huang, Shyh-Fang[2]
[1]Graduate Institute of Information and Decision Sciences, National Taipei University of Business, Taipei, Taiwan
[2]Department of Computer Science and Information Engineering, Minghsin University of Science and Technology, Taiwan

Abstract

Amid escalating global challenges such as climate change and resource depletion, the transition toward sustainable energy systems has become a critical priority. Solar energy, as a representative clean energy source, plays a pivotal role in reducing carbon emissions, enhancing energy security, and supporting corporate sustainability and environmental, social, and governance (ESG) performance. As the solar industry expands, competition among firms has intensified, making it increasingly important to balance technological innovation with market profitability to sustain long-term competitive advantage. While prior research has predominantly examined innovation and financial performance separately, limited attention has been given to their joint strategic implications, and few analytical tools simultaneously capture innovation, profitability, and sustainability, particularly within the solar energy sector. To address this gap, this study develops a Patent and Gross Margin Portfolio Management (PGMPM) model that integrates patent-based innovation indicators with profitability measures, providing a unified two-dimensional framework for evaluating firm-level strategic positioning. Using panel data from eight leading domestic and international solar energy firms over the period 2000–2024—drawing patent information from the LexisNexis database and financial data from the Datastream database—this study empirically examines the interaction between technological capability and financial performance. Firms are positioned within the portfolio matrix according to their patent counts and average gross margins, and their strategic profiles are analysed across the upstream, midstream, and downstream segments of the solar value chain. The results reveal substantial heterogeneity in firms' innovation–profitability profiles and demonstrate that firms achieving a balanced alignment between innovation and profitability exhibit stronger competitiveness and sustainability performance, whereas firms weak on both dimensions face significant strategic risk. The findings further highlight the strategic importance of patent portfolio management in enhancing ESG-related outcomes and long-term value creation. This study contributes to the literature by bridging innovation and profitability perspectives within a portfolio matrix framework and extending its application to the renewable energy sector. The proposed model offers actionable insights for strategic decision-making and supports the sustainable development of the solar energy industry, providing practical guidance for firms, policymakers, and industry stakeholders seeking to strengthen both technological and financial resilience.

Keywords: Solar cell industry analysis, Patent analysis, Corporate social responsibility, Environmental Social and Governance (ESG).

Paper type: Practical Paper

1. Introduction

Given the rapid depletion of fossil fuels, the development of renewable energy has become a strategic priority for governments and enterprises worldwide. According to the World Nuclear Association, recoverable oil reserves are estimated to last around 40 years, natural gas around 62 years, and coal around 216 years. As a result, countries have begun actively pursuing renewable sources to replace traditional fossil fuels, and studies increasingly examine green IT scheduling for renewable energy (Gu et al., 2018) and the economical operation of cloud data centres through energy trading with the power grid (Grange et al., 2018). Under the Renewable Energy Development Act (REDA), renewable energy includes solar, biomass, geothermal, ocean, wind, non-pumped hydropower, and energy recovered from waste. Among these, solar energy offers the greatest potential: the sun is a permanent, inexhaustible source of heat, and roughly 15 minutes of sunlight reaching the Earth's surface could, in principle, meet global energy demand for a full year.

Amid escalating challenges such as climate change and resource depletion, the solar energy industry has emerged as a key low-carbon sector. As the industry expands, however, competition has intensified, making it increasingly critical for firms to balance technological innovation with market profitability to sustain long-term competitive advantage. Existing studies have largely focused on single dimensions—using patent data to measure innovation capability, or financial indicators to evaluate performance—with limited effort to integrate these perspectives into a unified framework. With the growing importance of environmental, social, and

governance (ESG) considerations, firms are now expected to achieve not only technological and financial performance but also sustainable and socially responsible outcomes. Nevertheless, few analytical tools simultaneously capture innovation, profitability, and sustainability, particularly in the solar energy context.

To address this gap, this study proposes a two-dimensional portfolio matrix integrating innovation and profitability by combining patent-based indicators with gross-margin measures. Using the solar energy industry as the empirical context, the study analyses eight solar energy companies with the Patent and Gross Margin Portfolio Management (PGMPM) model, which combines firms' gross margins with their patent counts in a product-portfolio-management framework. The specific objectives are: (1) to analyse the current status of patent development in the global solar energy industry and examine the patent deployment strategies of major firms; (2) to apply the patent strategic matrix to assess firms' relative positions in technological innovation and market share; and (3) to explore how patent strategies influence corporate performance in environmental sustainability and green competitiveness, and to propose recommendations for future development.

This paper applies a correlational case-study design to eight listed domestic and foreign solar power firms, using the PGMPM business matrix to analyse the gross margins and patent counts of their products, to probe each firm's competitive advantage, and to derive policy recommendations. The implementation comprises: an introduction outlining research motives, objectives, and process; a literature discussion covering the solar cell industry and Ohmae's product-portfolio matrix; the proposal of the research framework and an overview of the domestic and international industry; the case study describing the solar cell industry and the PGMPM matrix calculations; and conclusions summarising implications for research management and strategic policy in the solar cell industry.

2. Literature Review

2.1 Patent Analysis

Citation-based patent analysis has been widely adopted (Albert et al., 1991; Breitzman & Thomas, 2002), using past patents as references for future ones, though it may underestimate newer patents that are cited less—a problem that worsens in fast-moving industries with short technology life-cycles. Analyses of solar energy patents have applied network and visual analysis (Kajikawa & Takeda, 2009), keyword-based approaches (Yoon & Park, 2004; Noh et al., 2015), high-tech trend analysis (Rafols & Meyer, 2010), and property–function analysis (Yoon & Kim, 2012), as well as technology-strategy studies of solar materials and cells (Torben et al., 2010; Wu & Mathews, 2012; Tseng et al., 2011).

In innovation research, patents are commonly used to measure firms' technological capabilities and knowledge accumulation, capturing not only innovation outputs but also technological quality, knowledge impact, and competitive positioning. Scholars increasingly argue that patent evaluation should extend beyond financial valuation to incorporate non-financial dimensions such as technological relevance and strategic implications, and multidimensional frameworks have been proposed to strengthen patent analysis in strategic decision-making (Lawson-Drackey, 2012). Patent-based analyses also distinguish between component and architectural knowledge, offering insight into the breadth and depth of technological development (Meunier & Bellais, 2019), and enable the identification of technological trajectories and lifecycle stages for forecasting and planning (Lee et al., 2024; Coccia & Roshani, 2025).

At the firm level, technological capabilities proxied by patents can enhance investment efficiency, improve resource allocation, and reduce information asymmetry (Li et al., 2025), and patent texts help explain firm market value and predict future earnings (Zheng, 2025). The relationship is not always linear, however; innovation may introduce uncertainty and, under certain conditions, even lead to overvaluation (Doruk, 2025). With advances in artificial intelligence, machine learning, semantic analysis, and text mining are increasingly applied to identify technological topics and innovation patterns—for example, vector-space and semantic methods that assess technological relevance (Huang & Lin, 2018), and multi-source approaches combining patents with web news to forecast emerging trends (Li et al., 2025).

In the solar energy sector—technology-intensive and policy-driven—innovation depends heavily on continuous R&D and patent deployment. Patent data have been used to analyse technological trends and knowledge flows in solar technologies (Chen & Pham, 2014), while government policy, technological capability, and market demand are considered key drivers of industry development (Wang & Wang, 2020), and global competition requires firms to balance innovation investment with cost efficiency (Guo et al., 2025). Meanwhile, ESG factors have emerged as critical determinants of firm performance: ESG affects stock returns, risk, and investment decisions (Teja & Liu, 2024); integrating ESG into predictive models improves accuracy (Park et al., 2024); ESG has evolved from a compliance requirement into a source of competitive advantage (Xing et al.,

2025); and ESG controversies can reduce firm value, moderated by governance mechanisms such as board diversity and independence (Brinette et al., 2024). Overall, most studies address a single dimension and rarely integrate innovation, profitability, and sustainability into one framework—the gap this study addresses through a two-dimensional portfolio matrix applied to the global solar energy industry.

2.2 Product Portfolio Management

Kenichi Ohmae's Product Portfolio Management (PPM) business matrix combines the Boston Consulting Group (BCG) matrix and the GE–McKinsey matrix. The BCG matrix, designed by Bruce Henderson in 1970, helps firms assess an existing product line and allocate cash efficiently (Hax & Majluf, 1983; Hill & Jones, 1995); its horizontal axis represents relative market share and its vertical axis market growth rate, yielding four product types—Question Marks, Stars, Cash Cows, and Dogs. In the PPM career matrix, the horizontal axis represents company strength (divided into three levels) and the vertical axis the attractiveness of the external industry (also three levels). The company-strength and external-environment variables are summarised in Table 1.

Table 1: Company strength variables and indicators of external environment variables

Company strength	External environment
1. Competitiveness: market share 2. Profits: profit margin, cash flow 3. Resources: financial and human resources, raw materials, downstream management 4. Technical: patented and unique technology, quality and design 5. Marketing: brand, regular customers, promotion, marketing dominance	1. Surrounding environment: social, environmental, legal, political, organizational, industry-intensive 2. Profits: industry profit tendency and profitability 3. Competition status: occupancy, capacity, entry difficulty, vertical integration, service and price elasticity 4. Market: size, growth rate, stability, level of detail

Source: Kenichi Ohmae, corporate consultant.

Beyond power generation, solar cells are widely used in consumer, communication, and signalling products (Mary, 1991; Pilkington et al., 2002). The commercial development of solar energy is regarded as an important trend internationally: the European Union, Japan, and the United States have prioritised solar and other renewables as a safe post-2030 energy supply, and solar power is projected to exceed 10% of world electricity supply by 2030 and 20% by 2050. Most major countries promote renewables through tax concessions, equipment subsidies, and preferential feed-in tariffs, using public funds to accelerate R&D and build enabling infrastructure. Although the 2008 US subprime crisis and the subsequent European debt crisis depressed markets and disrupted the industry, the global solar market has continued to advance and grow over the long term.

3. Methodology

This study uses Patent and Gross Margin Portfolio Management (PGMPM) matrix analysis to investigate the solar cell industry and profile the case firms. The PGMPM business matrix combines patent analysis with product portfolio management to evaluate the priority of patent value and firm performance, identify risks and opportunities, prioritise products, and optimise the portfolio, thereby supporting strategies to achieve target revenue and profitability. The PGMPM algorithm takes as input (i) patent sets and (ii) the gross-margin portfolio, and outputs a characteristic rule set generalised from the input, through four steps: (1) collect the patent-set data; (2) perform the PGMPM model—beginning from the basic patent sets, for each attribute in the gross-margin portfolio, selectively generalise attributes and merge identical rules; (3) describe the solar cell industry and compute the PGMPM business matrix; and (4) summarise the conclusions and recommendations for strategic policy in the solar cell industry.

The PGMPM framework provides the foundation for a case study of the solar energy industry. The case firms are listed companies whose solar cell patent information is registered in the LexisNexis database, while their financial statements are obtained from the Datastream database. The eight firms examined are Sino-American Silicon, NEO Solar Power, KYOCERA, Evergreen Solar, BP Solar, Wacker Chemie, Q-Cells, and KANEKA. Sino-American Silicon and NEO Solar Power are professional wafer-material suppliers focused on solar plant development and production; KYOCERA extended from high-precision ceramics into solar cells as a diversified high-tech enterprise; Evergreen Solar is a US solar cell manufacturer; BP Solar is a leading global solar company; Wacker Chemie operates in organic silicon, semiconductors, and chemical polymerisation; Q-Cells is a leading

solar panel manufacturer; and KANEKA mainly produces chemicals, functional resins, and electronic components. Profiles of the firms are summarised in Table 2.

Table 2: Information on the various solar companies

Company	Profile
Sino-American Silicon	Professional wafer-material supplier.
NEO Solar Power	Solar energy R&D and production plant, established in 2005.
KYOCERA	Former high-precision ceramic factory, now extended into solar cells as a diversified high-tech enterprise.
Evergreen Solar	US solar cell manufacturer.
BP Solar	One of the world's leading solar companies.
Wacker Chemie	Global chemical company operating in organic silicon, semiconductors, and chemical polymerisation.
Q-Cells	Leading solar panel manufacturer.
KANEKA	Mainly produces chemicals (PVC), functional resin, and electronic components.

4. Solar Power Plant PGMPM Business Matrix Analysis

The solar cell industry can be divided into upstream, midstream, and downstream segments. The upstream processes materials: after obtaining silicon from suppliers, a crystal-growth furnace produces a silicon ingot, which a precision cutting machine slices into wafers. The midstream focuses on solar cell components (PV cells), and the downstream on product applications, including PV modules and assembled solar systems. The eight firms were analysed with the PGMPM matrix, where the horizontal axis represents company strength (proxied by total patent count from LexisNexis) and the vertical axis the external environment (proxied by the average gross-margin ranking over 2000–2024, from Datastream). Firm rankings are shown in Table 3.

Table 3: Companies' gross margins and number of patents

Company	Avg. gross margin	Margin rank	Patents (upper/mid/lower)	Patent rank (total)
Sino-American Silicon	26.37%	2	0 / 22 / 19	1 (41)
NEO Solar Power	27.62%	1	8 / 15 / 8	2 (31)
KYOCERA	15.26%	6	0 / 21 / 4	3 (25)
Evergreen Solar	-40.99%	8	11 / 7 / 6	4 (24)
BP Solar	20.86%	4	14 / 6 / 1	5 (21)
Wacker Chemie	15.85%	5	1 / 7 / 6	6 (14)
Q-Cells	26.33%	3	13 / 0 / 0	7 (13)
KANEKA	10.28%	7	0 / 7 / 1	8 (8)

Plotting each firm by its total patent count and average gross margin yields the solar energy PGMPM business matrix shown in Figure 1. NEO Solar Power and Sino-American Silicon combine relatively high patent counts with strong gross margins, placing them in a favourable position; their strategy is to adhere to advantages, expand industrial scale, and selectively increase investment in response to future trends. Sino-American Silicon, as a professional foundry, relies mainly on order volume, so the pressure for new technology is less acute; actively developing new technology would enhance its competitiveness and reduce order dependence. Wacker Chemie shows good industry attractiveness but a low patent count; as an upstream raw-material supplier, its investment plan—spending EUR 200 million to expand polysilicon production and fund new R&D, targeting about 22,000 metric tons (roughly 13% of global production)—can be read as an attempt to dominate the upstream.

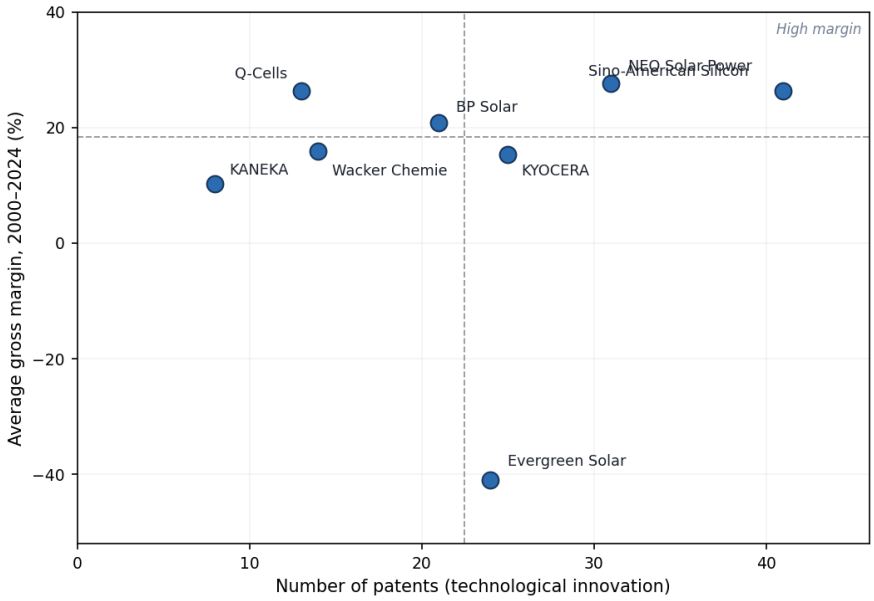


Figure 1: The solar energy PGMPM business matrix

NEO Solar Power, BP Solar, and Q-Cells rank in the middle for patent count. NEO Solar Power's gross margin is below the eight-firm average; consistent with the matrix, its feasible strategies are to eliminate risk and increase profits. Established at the end of 2005, its market share is still limited, suggesting a need for closer cooperation with other manufacturers. BP Solar and Q-Cells sit in the middle of the matrix, making them the most flexible and able to adjust to circumstances; the BP oil-spill incident, which cost the company more than USD 400 billion, illustrates such exposure, and although its revenue has recovered it has not returned to prior levels.

Evergreen Solar has recorded significant losses over the past decade; following a net loss of about USD 12.5 million, it announced plant shutdowns and layoffs and later warned of a funding shortage driven by declining orders, raising bankruptcy risk. Its position corresponds to revenue maximisation under the matrix—where cash cows can no longer generate profit—implying it is likely to divest plants or exit. KYOCERA and KANEKA, by contrast, occupy strong positions: KYOCERA began developing solar cells in 1975, reaching single-crystal cell efficiency of 19.5% and substantial annual output, and expanded production capacity internationally, while KANEKA invested heavily in Hyogo Prefecture targeting 1 GW capacity. Both firms diversified into solar from other industries—precision ceramics and chemicals respectively—reflecting a standard growth-investment strategy anchored on scale advantage.

Value-chain analysis of the patent statistics shows that Sino-American Silicon and Wacker Chemie focus upstream: both aim to improve silicon purity, but Sino-American Silicon emphasises equipment efficiency (crystal-pulling and growth furnaces) to produce high-quality silicon rods, while Wacker Chemie focuses on purifying and processing monocrystalline silicon, which is more complex but yields higher generation efficiency. Evergreen Solar and KYOCERA hold upstream patents in silicon wafers and ingots. The remaining firms concentrate on midstream and downstream stages—PV components and module packaging: KANEKA develops next-generation thin-film cell materials; KYOCERA, Evergreen Solar, and NEO Solar Power hold electrode-improvement patents; BP Solar develops amorphous-silicon and dye-sensitised cells; and Q-Cells holds back-contact cell technology that increases the light-absorbing area for higher efficiency than conventional front-electrode designs. KYOCERA and BP Solar also hold chemical-vapour-deposition patents for anti-reflective and reflective coatings that improve sunlight absorption. Downstream, KANEKA offers the most complete capability from module to package testing, while KYOCERA holds a distinctive patent for solar-tracking devices. Although these data do not by themselves prove that Taiwan's R&D lags other countries, they indicate that domestic capabilities still need strengthening through systematic patent analysis to inform strategy, market direction, and talent development.

5. Conclusions and Suggestions

5.1 Solar Cell Energy Industry Strategy Management

As the world promotes energy-saving technology, solar energy has become a major player in the green energy industry. Taiwan's well-developed foundry sector remains an important revenue source for many small and medium enterprises, but the industry should transform toward a knowledge-based model, which requires cooperation among production, government, academia, and research to train technical personnel—supported, for example, by the Ministry of Education's green-technology talent-cultivation programme. Technology and marketing are further levers: the cost of producing 1 W of capacity remains high even with subsidies, and cells must last more than 20 years to balance costs. Upgrading technology—through transfer of foreign technology or domestic development of high-conversion cells—and changing marketing strategies to increase installations in public facilities and residential roofs are therefore essential. In the future, solar products must be diversified, reach high value, and promote environmental protection; improving conversion efficiency and subsidising production capacity remain the most important issues for both the industry and government.

5.2 Management Implications and Future Research

This study makes two contributions. First, it proposes the PGMPM model, which combines patent analysis with gross-margin portfolio management to obtain, integrate, accumulate, transfer, and innovate an industry's knowledge value. Second, it applies the model to the solar energy industry to integrate the major value attributes of the sector. In the internal-environment analysis, the industry maintains good policy convergence, but industry consolidation and marketing still have room for improvement; achieving cluster efficiency requires a strong value chain that combines government and industry and emphasises the layout of technology patents. In the external-environment analysis, Europe, the United States, and Japan each present clear solar strategies and development visions; Taiwan now offers renewable-energy subsidies, feed-in tariffs, equipment subsidies, low-interest loans, and tax relief, signalling policy commitment, but the industry still lacks horizontal and vertical integration.

Based on the analysis, recommendations for the solar energy industry fall into six directions: (1) government and private enterprises should jointly create green-energy employment, integrate industry and market planning with policy, expand financing channels, establish an industrial development investment fund, and build a national quality-supervision mechanism; (2) strengthen patents, technology research, and forecasting in key scientific fields to shape manufacturers' advantages and master dominant specifications; (3) overcome reliance on rare metals such as germanium and improve conversion efficiency, with academia and industry cooperating to reduce R&D costs; (4) use lens-based light concentration to improve generation efficiency and reduce cost while enhancing device capacity and durability; (5) promote cluster efficiency and a sound value chain that combines upstream, midstream, and downstream strengths, supported by new energy resources, geographic information systems, and personnel training; and (6) focus on policy integration, industrial integration, and marketing, learning from and cooperating with Europe, the United States, and Japan to pursue multiple strategic layouts and global promotion.

6 Acknowledgements

The authors thank the colleagues and institutions that supported the data collection and analysis for this study.

7 Ethics declaration

This study relied exclusively on publicly available secondary data—patent records from the LexisNexis database and financial statements from the Datastream database. No human participants, personal data, or sensitive information were involved. Accordingly, formal ethics committee clearance was not required for this research.

8 AI declaration

All research design, data collection, analysis, interpretation, and conclusions are the authors' own, and the authors reviewed and verified all content to ensure its accuracy and integrity.

References

Albert, M., Avery, D., Narin, F. and McAllister, P. (1991) “Direct validation of citation counts as indicators of industrially important patents”, Research Policy, Vol. 20, No. 3, pp 251–259.

Breitzman, A. and Thomas, P. (2002) “Using patent citation analysis to target/value M&A candidates”, Research-Technology Management, Vol. 45, No. 5, pp 28–36.

Brinette, S., Sonmez, F.D. and Tournus, P.S. (2024) “ESG controversies and firm value: moderating role of board gender diversity and board independence”, IEEE Transactions on Engineering Management, Vol. 71, pp 4298–4307.

Chen, J.K.C. and Pham, V.K. (2014) “A study on knowledge flows of dye-sensitized solar cells' patent”, Foresight, Vol. 16, No. 3, pp 229–249.

Coccia, M. and Roshani, S. (2025) “Path-breaking directions in quantum computing technology: a patent analysis with multiple techniques”, Journal of the Knowledge Economy, Vol. 16, No. 1, pp 4991–5024.

Doruk, O.T. (2025) “The link between COVID vaccine patents and fundamental uncertainty in financial markets: a post Keynesian analysis”, Journal of Post Keynesian Economics, Vol. 48, No. 2, pp 358–373.

Grange, L., Costa, G.D. and Stolf, P. (2018) “Green IT scheduling for data centers powered with renewable energy”, Future Generation Computer Systems, Vol. 86, pp 99–120.

Gu, C., Fan, L. and Wu, W. et al. (2018) “Greening cloud data centers economically by energy trading with the power grid”, Future Generation Computer Systems, Vol. 78, No. 1, pp 89–101.

Guo, K., Zhu, H. and Yu, F. et al. (2025) “It's not overcapacity, but a matter of global imbalances”, China & World Economy, Vol. 33, No. 2, pp 138–170.

Hax, A.C. and Majluf, N.S. (1983) “The use of the growth-share matrix in strategic planning”, Interfaces, Vol. 13, No. 1, pp 46–60.

Hill, C. and Jones, G. (1995) Strategic Management Theory: An Integrated Approach, Houghton Mifflin, Boston.

Huang, Y. and Lin, C. (2018) “Innovation topic analysis using vector space model: a patent-based approach”, IEEE Conference Proceedings.

Kajikawa, Y. and Takeda, Y. (2009) “Citation network analysis of organic LEDs”, Technological Forecasting and Social Change, Vol. 76, No. 8, pp 1115–1123.

Lawson-Drackey, S. (2012) “Méthode d'évaluation des brevets dans une perspective multidimensionnelle”, Économies et Sociétés, Vol. 46, No. 1, pp 97–123.

Lee, Y.-J., Han, Y.J. and Kim, S.-S. et al. (2024) “Identifying technology opportunities and taxonomy for railway static inverters using patent data analytics”, IEEE Access, Vol. 12, pp 17389–17403.

Li, X., Wang, S. and Dai, P. et al. (2025) “Big data technology and corporate investment efficiency: evidence from firm-level patent data”, Applied Economics, Vol. 57, No. 41, pp 6572–6591.

Li, Y., Wang, H. and Zhao, Q. (2025) “Emerging technologies using patent analysis and web news data mining: the case of perovskite solar cell technology”, IEEE Transactions on Engineering Management.

Mary, E.M. (1991) “Using patent data for technology analysis and planning”, Research-Technology Management, pp 43–49.

Meunier, F.X. and Bellais, R. (2019) “Technical systems and cross-sector knowledge diffusion: an illustration with drones”, Technology Analysis and Strategic Management, Vol. 31, No. 4, pp 433–446.

Noh, H., Jo, Y. and Lee, S. (2015) “Keyword selection and processing strategy for applying text mining to patent analysis”, Expert Systems with Applications, Vol. 42, No. 9, pp 4348–4360.

Park, J., Na, H.J. and Kim, H. (2024) “Development of a success prediction model for crowdfunding based on machine learning reflecting ESG information”, IEEE Access, Vol. 12, pp 197275–197289.

Pilkington, A., Romano, D. and Omid, T. (2002) “The electric vehicle: patent data as indicators of technological development”, World Patent Information, Vol. 24, No. 1, pp 5–12.

Rafols, I. and Meyer, M. (2010) “Diversity and network coherence as indicators of interdisciplinarity”, Scientometrics, Vol. 82, No. 2, pp 263–287.

Teja, K.R. and Liu, C.-M. (2024) “ESG investing: a statistically valid approach to data-driven decision making and the impact of ESG factors on stock returns and risk”, IEEE Access, Vol. 12, pp 69434–69444.

Torben, D., Nielsen, T.D. and Cruickshank, S.F. et al. (2010) “Business, market, and intellectual property analysis of polymer solar cells”, Solar Energy Materials & Solar Cells, Vol. 94, pp 1553–1571.

Tseng, F.M., Hsieh, C.H. and Peng, Y.N. et al. (2011) “Using patent data to analyze trends and technological strategies of amorphous silicon thin-film solar cells”, Technological Forecasting and Social Change, Vol. 78, pp 332–345.

Wang, M.L. and Wang, K.W. (2020) “A case study of the prospects of solar industry players”, International Journal of Economics and Research, Vol. 11, No. 3, pp 89–110.

Wu, C.Y. and Mathews, J.A. (2012) “Knowledge flows in the solar photovoltaic industry”, Research Policy, Vol. 41, No. 3, pp 524–540.

Xing, Z., Huang, J. and Fang, D. (2025) “From compliance to competitiveness: unpacking the impact of ESG performance on strategic innovation and market dynamics”, IEEE Transactions on Engineering Management, Vol. 72, pp 1297–1319.

Yoon, B. and Park, Y. (2004) “A text-mining-based patent network”, Journal of High Technology Management Research, Vol. 15, No. 1, pp 37–50.

Yoon, J. and Kim, K. (2012) “Patent network analysis for strategic R&D planning”, Expert Systems with Applications, Vol. 39, No. 9, pp 7709–7717.

Zheng, Y. (2025) “Can investors learn from patent documents? Evidence from textual analysis”, Contemporary Accounting Research, Vol. 42, No. 2, pp 1331–1358.

Legal Risk Management of Business Owners from the Perspective of Choice-of-Law Rules: A Research Focus On the Impact of Environmental Change on Food Safety

Chang, Che-Wei, Hung, Yu-Chun, Tsai, Pei-Fen
Asia University, Taiwan (R.O.C.)

Abstract

Different choice-of-law rules give rise to vastly different corporate liabilities. This article investigates the management of liability risks and cost controls for multinational enterprises (MNEs) or International Trading Company, specifically addressing how such entities can preemptively mitigate the disparate legal exposures arising from the divergence of conflict-of-law rules. The core of this research focuses on tortious liability for damages stemming from food safety and hygiene crises precipitated by environmental transformations, examining the strategic navigation of applicable legal frameworks to ensure institutional stability.

The Rome II Regulation establishes a unified framework for choice-of-law rules regarding non-contractual obligations (torts) across all EU Member States. Specifically, Article 7 stipulates the choice-of-law regime for environmental damage, which consequently subjects enterprises to varying liability risks depending on the applicable governing law. Regarding the legal framework in Taiwan, the relevant choice-of-law rules are stipulated in Articles 26 of the Act Governing the Application of Laws in Civil Matters Involving Foreign Elements. Unlike the EU framework, Taiwanese law currently lacks a specialized provision for environmental torts, thereby mandating that such cases be governed by general tortious choice-of-law rules.

Due to these diverging protective purposes and developmental trajectories, the stringency of these frameworks varies significantly. In specific food safety cases, the application of EU law under choice-of-law rules substantially increases the probability of successful claims by victims. Consequently, when EU law is applicable, business owners must engage in quality control standards commensurately elevated to manage legal obligations effectively.

If managerial duties are governed by Taiwanese law, which aligns with United Nations standards, an additive may be permitted in the absence of conclusive evidence of harm. That is, as long as the product does not pose an 'unreasonable' or 'unacceptable' risk to public health under reasonably foreseeable and normal conditions of use, the addition of such a substance will generally not be unlawful.

Multinational Corporation or International Trading Company must anticipate applicable choice-of-law outcomes. When EU law applies, enterprises cannot rely on permissive local standards; they must proactively elevate internal quality controls and supply chain standards to satisfy strict precautionary requirements and mitigate potential compensation claims.

Herefore, multinational enterprises or international trading companies must anticipate applicable choice-of-law outcomes to serve as control standards for internal quality assurance and supply chain management.

Keywords: Conflict-Of-Law Rules, Legal Risk Management, Environmental Change, Food Safety, Tortious Liability

Paper type: an Academic Research Paper

1 Research Subjects and Scope

This study examines how enterprises can proactively mitigate legal risks of tort liability for damages when environmental pollution, or environmental changes caused by climate change, give rise to food safety issues. [1] Because tortious conduct encompasses a wide range of circumstances, and to facilitate analysis of the relationship between choice-of-law rules and the liabilities of transnational enterprises, the study confines its scope to the impact of environmental and climatic changes on food hygiene and safety.

Corporate rights and obligations concerning food hygiene and safety can be categorized primarily under three regimes: EU standards, United Nations standards, and US standards. Taiwan tends to align more closely with UN and US frameworks. This article therefore delineates the principal differences among these three regimes and examines how the liability of corporate actors shifts when choice-of-law rules select a particular national or supranational governing law and its corresponding substantive rules are applied.

2 Food Additives and Case Studies

Suppose a European tourist visiting the United States purchases and consumes a food product that was manufactured in Taiwan and exported to the U.S. After ingestion, the tourist experiences itching and

[1] *For reasons related to the study's theoretical framework and word-count constraints, this research is limited to tort liability and does not address damages arising from contractual obligations.*

gastrointestinal inflammation. Examination of the product's ingredients raises suspicion that the adverse reaction was caused by the artificial colorant titanium dioxide (Titanium Dioxide, E171/INS 171). The tourist then brings a tort claim seeking damages, including medical expenses, compensation for emotional distress, transportation costs for medical consultations, and lost earnings due to time off work for treatment.

Whether titanium dioxide (Titanium Dioxide, E171/INS 171) may be lawfully used as a food additive is currently one of the most significant points of international disagreement. In 2022 the EU concluded that titanium dioxide could not be excluded as causing DNA damage, identifying a potential genotoxic risk and expressing concern that particle accumulation might induce genotoxic effects or chronic intestinal inflammation, and consequently enacted a comprehensive ban[1]. By contrast, the Joint FAO/WHO Expert Committee on Food Additives (JECFA) assessed that the available data remain insufficient to demonstrate an immediate hazard to humans; accordingly, the Codex Alimentarius Commission (CAC) lists titanium dioxide as an authorized white pigment in the General Standard for Food Additives (GSFA). In Taiwan, under the Ministry of Health and Welfare—Taiwan Food and Drug Administration (TFDA)'s Standards for the Scope of Use, Maximum Limits and Specifications of Food Additives, titanium dioxide is classified as a Category IX colorant and may be used in various foodstuffs as needed within appropriate limits; however, it is expressly prohibited from use in certain unprocessed foods, including fresh meats, fresh seafood, fresh legumes, fresh vegetables, fresh fruits, miso, soy sauce, kelp, nori, and tea[2]. If meat becomes discolored and turns gray, the addition of titanium dioxide can make it appear redder; likewise, tofu skin or surimi products may incorporate titanium dioxide to mask inferior raw materials. Soy sauce's color should derive from fermentation and caramelization, and tea from the natural hue of the leaves; permitting titanium dioxide would undermine these traditional color characteristics, risk misleading consumers, and could result in purchases of stale or poor-quality foods. Accordingly, on 23 March 2026 the authorities issued a draft "Analytical Methods and Specifications for Titanium Dioxide as a Food Additive," [3]indicating governmental efforts to strengthen monitoring of the substance's purity and composition, although this measure does not amount to an outright ban.

The EFSA assessment indicates that occurrences of typical food-allergic reactions to titanium dioxide (e.g., urticaria or respiratory difficulties) are exceedingly rare[4]. A 2024 study investigating the impact of titanium dioxide on the intestinal barrier reported that long-term ingestion of micrometer- or nanometer-sized titanium dioxide may damage the intestinal mucosa. When the gut barrier is compromised (so-called "leaky gut")[5], systemic inflammatory responses may be heightened, potentially exacerbating preexisting conditions; although clinical case reports are limited, their scarcity does not imply absence of risk.

3 Liability of Product Manufacturers

3.1 Legal Nature and Classification

In the case described above, the sale arises from a contract between the retail merchant that sold the product and the consumer; the product manufacturer has no contractual relationship with the buyer. Accordingly, any recovery sought against the manufacturer should be grounded in tort law, and the legal characterization of the claim is one of manufacturer liability in tort.

The product was manufactured by a Taiwanese company whose principal is a national of the Republic of China. The food, produced in Taiwan, was imported into and complied with U.S. food regulations. A European consumer who returns to Europe and then consumes the Taiwan-made product purchased in the United States and subsequently seeks medical treatment may have been harmed for various reasons: a defect intrinsic to the product; deterioration during prolonged transport or storage after leaving the U.S. and en route to Europe due to changes in climate and environment; or a post-consumption interaction between the product and other foods

---

[1] FAF Panel EFSA Panel on Food Additives and Flavourings (FAF) (2021), Scientific Opinion on the safety assessment of titanium dioxide (E171) as a food additive, EFSA Journal, May, pp.1,9,73-87, https://doi.org/10.2903/j.efsa.2021.6585 (as available on 18.5.2026)

[2] Standards for the Scope of Use, Maximum Limits, and Specifications of Food Additives (amended on 7 May 2025 [R.O.C. Year 114]), Annex 1: Scope of Use and Maximum Levels of Food Additives – Category IX: Colorants. Available at: chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://db.lawbank.com.tw/GetFile.ashx?pfid=0000297834 (as available on 18.5.2026)

[3] Taiwan Food and Drug Administration (TFDA), Ministry of Health and Welfare, preannounced a draft revision of the "Analytical Methods and Specifications for Titanium Dioxide as a Food Additive" on 23 March 2026. https://www.fda.gov.tw/tc/newsContent.aspx?cid=5072&id=31461 (as available on 18.5.2026)

[4] European Commission (2016), Question number:EFSA-Q-2011-00348, Re-evaluation of titanium dioxide (E 171) as a food additive-- EFSA Panel on Food Additives and Nutrient Sources added to Food (ANS), First published:14 September 2016, EFSA Journal, Vol. 14, No. 9, pp.57-68, See https://efsa.onlinelibrary.wiley.com/doi/10.2903/j.efsa.2016.4545 (as available on 18.5.2026)

[5] "The disruption of the intestinal epithelial barrier allows for the translocation of pro-inflammatory cytokines into the systemic circulation... which can potentially exacerbate pre-existing skin conditions such as chronic urticaria or atopic dermatitis by lowering the degranulation threshold of skin mast cells." Niqhat Fatma Rizvi, Anoop Kumar, Rashmi Chandra, and Shoebul Haque (2024), Synthetic Food Colourants: Applications, Regulatory Perspectives, and Potential Health Implications, World Journal of Biology Pharmacy and Health Sciences (WJBPHS), Vol.17, No. 1, January, p.63. See DOI: 10.30574/wjbphs.2024.17.1.0068 (as available on 18.5.2026)

producing a chain reaction in the body despite the product itself remaining unspoiled. These distinct causal scenarios give rise to different patterns of corporate liability, and litigation brought in European courts versus courts of the Republic of China may produce divergent outcomes (discussed further below).

In cross-border disputes, the initial decision is which court will hear the case—i.e., the choice of international jurisdiction. Once jurisdiction is established, the next step is to determine the applicable law. In tort cases, and to avoid divergences among national substantive rules, the EU adopted unified choice-of-law rules for non-contractual obligations known as Rome II (Regulation (EC) No. 864/2007, commonly referred to as Rome II).

If the plaintiff files suit in Taiwan on the basis that the defendant manufacturer is Taiwanese, the choice-of-law instrument applied by Taiwanese courts is the Act Governing the Choice of Law in Civil Matters Involving Foreign Elements.

3.2 Corporate Liability under the EU Choice-of-Law Rules

Under the Brussels I Recast (Regulation (EU) No 1215/2012) [1], the courts of the place where the tort was committed—including both the place of the harmful event and the place where the damage occurred[2]—are competent to exercise international jurisdiction.

For the benefit of Taiwanese business owners in understanding that corporate liability in transnational cases depends on choice-of-law rules—and that different rules may yield divergent liabilities—this paper outlines the choice-of-law provisions most commonly applied to manufacturers' torts: Article 5 of Rome II for product-related non-contractual obligations, and Article 7 of Rome II for torts arising from changes in climate or the environment.

3.2.1 Corporate Liability for Environmental Damage under the Choice-of-Law Rule of Article 7 of Rome II

Article 7 (Environmental damage) of Rome II: "The law applicable to a non-contractual obligation arising out of environmental damage or damage sustained by persons or property as a result of such damage shall be the law determined pursuant to Article 4(1), unless the person seeking compensation for damage chooses to base his or her claim on the law of the country in which the event giving rise to the damage occurred." If, pursuant to Article 4(1), the law of the country where the damage occurred is to be applied, then the governing law is determined solely by the place of the damage regardless of the country in which the factual cause of the damage arose or where any consequential or indirect effects occurred. Applying this to the present case, EU law would govern. Because the inclusion of titanium dioxide in the product contravenes EU food regulations, the manufacturer should be held liable for the harm suffered by the consumer, an outcome that is unfavorable to the manufacturer.

Article 7 of Rome II also provides an alternative choice-of-law criterion: the law of the country in which the event giving rise to the damage occurred. Accordingly, the defendant may contend that the applicable law in this case should be that of the United States because the product was originally acquired there and the causal event giving rise to the damage occurred there; thus U.S. law, rather than EU law, should govern. Under U.S. law the use of titanium dioxide is lawful, so the mere inclusion of titanium dioxide would not, per se, give rise to legal liability. Liability would instead depend on whether environmental or climatic changes led to product deterioration that caused the consumer's illness, an issue requiring factual and expert inquiry. Moreover, under both the Second[3] and Third[4] Restatements, the predicate for liability is a product defect, which is understood to include manufacturing defects, design defects, and defects arising from inadequate instructions or warnings.[5] Accordingly, if a manufacturer or commercial operator can demonstrate that the product was not defective, the requisite elements for tort liability are lacking and no compensation is owing to the plaintiff. The defendant may further argue that, if the harm resulted from improper storage during long-distance transportation, customs detention, or the plaintiff's own mishandling, the risk should be borne by the plaintiff.

---

[1] Regulation (EU) No 1215/2012 of the European Parliament and of the Council of 12 December 2012 on jurisdiction and the recognition and enforcement of judgments in civil and commercial matters (Brussels I Recast), OJ L 351/1, 2012.

[2] The three principal provisions of the Brussels I Regulation are: (1) the defendant's domicile rule (Article 4); (2) the tort venue rule, encompassing both the place of the wrongful act and the place where the damage occurred (Article 7); and (3) consumer-protection jurisdiction (Articles 17–19).

[3] Article 402A of Restatement (Second) of Torts: "Special Liability of Seller of Product for Physical Harm to User or Consumer (1) One who sells any product in a defective condition unreasonably dangerous to the user or consumer or to his property is subject to liability for physical harm thereby caused to the ultimate user or consumer, or to his property,....."

[4] Article 1 of Restatement (Third) of Torts: Products Liability : "Liability of Commercial Product Seller for Harm Caused by Defective Products-One engaged in the business of selling or otherwise distributing products who sells or distributes a defective product is subject to liability for harm to persons or property caused by the defect."

[5] Article 2 of Restatement (Third) of Torts: Products Liability: "Categories of Product Defect -A product is defective when, at the time of sale or distribution, it contains a manufacturing defect, is defective in design, or is defective because of inadequate instructions or warnings."

3.3.2 Corporate Liability under the Choice-of-Law Rule of Article 5(1) main paragraph of Rome II

It is important to note the Article 5(1) textual proviso—"subject to Article 4(2)"—which reflects that Article 5(1) applies where the liable party and the victim are not both habitually resident in the same country at the time the damage occurs. If the manufacturer or liable party is not domiciled or habitually resident in an EU Member State, the assumptions underpinning Article 5(1) main paragraph may not be met, and the special rule set out in the proviso to Article 5(1) must be applied instead.

3.3.3 Corporate Liability under the Proviso to Article 5(1) of Rome II

According to the proviso to Article 5i of Rome II :"However, the law applicable shall be the law of the country in which the person claimed to be liable is habitually resident if he or she could not reasonably foresee the marketing of the product, or a product of the same type, in the country the law of which is applicable under (a), (b) or (c)." In other words, If the producer could reasonably have foreseen that the product might be brought into use in the EU, and the producer's habitual residence is in an EU Member State, the law of that Member State will apply and the producer will be liable for adverse health effects resulting from the addition.

The capacity to foresee and the fact of having foreseen must be clearly distinguished. Legal responsibility attaches to the capacity to foresee: if a defendant lacked that capacity, it should not be reproached; if the defendant possessed that capacity, then irrespective of whether actual foresight occurred, the defendant is obliged to implement measures commensurate with that foresight, and failure to act gives rise to liability. Whether the capacity to foresee existed is a matter for judicial appraisal, and accordingly the producer faces a risk of being ordered to pay damages.

3.3.4 Corporate Liability under the Choice-of-Law Rule of Article 5(2) of Rome II

According to Article 5 of Rome II: " 2.Where it is clear from all the circumstances of the case that the tort/delict is manifestly more closely connected with a country other than that indicated in paragraph 1, the law of that other country shall apply. A manifestly closer connection with another country might be based in particular on a pre-existing relationship between the parties, such as a contract, that is closely connected with the tort/delict in question.", based on a "manifestly closer connection" with another country, a company can argue that its products were manufactured in Taiwan and sold in the United States, thus establishing a close connection with Taiwan or the United States. The EU connection is merely an isolated incident, therefore Taiwanese or US law, rather than EU law, should apply.

In this case, EU consumers brought the goods back from the US and the product spoiled due to local climate or environment, resulting in health problems, then the consumer's actions constitute an intervention or processing that caused product damage. Even if the court finds the company liable, the manufacturer can argue that the plaintiff's negligence should be offset when calculating damages to reduce its liability.

3.3 Corporate Legal Liability under Taiwan's Choice-of-Law Rules

3.3.1 Corporate Liability under the Main Paragraph of Article 26 of the Act Governing the Choice of Law in Civil Matters Involving Foreign Elements

If this case is litigated in Taiwan, the Taiwanese court will apply Article 26 of Act Governing the Choice of Law in Civil Matters Involving Foreign Elements: "Where an injury has resulted from an ordinary use or consumption of an article of commerce, the legal relationship between the injured person and the manufacturer is governed by the national law of the latter. However, where the manufacturer has agreed in advance or where the manufacturer could have foreseen that the article would be sold in a place whose law is one of the three mentioned below, the law of that place is applied, if the injured person chooses that law as the applicable law: a. The law of the place of injury; b. The law of the place where the injured person purchased the article; and c. The national law of the injured person." Under that provision, where the applicable law is "the law of the country of the manufacturer," if the manufacturer is a national of the Republic of China (Taiwan) Taiwanese civil law and related statutes apply. The addition of titanium dioxide would therefore be lawful and not by itself actionable.

However, if consumption of the food causes adverse health effects, liability may arise pursuant to Article 191-1 of the Civil Code[1], a manufacturer or commercial enterprise should be held liable for damages unless it

[1] *Article 191-1 of the Civil Code : "The manufacturer is liable for the injury to another arising from the common use or consumption of his merchandise, unless there is no defectiveness in the production, manufacture, process, or design of the merchandise, or the injury is not caused by the defectiveness, or the manufacturer has exercised reasonable care to prevent the injury.*

can demonstrate that the product's production, manufacture, processing, or design was free of defect, that the harm did not result from such a defect, or that it exercised due care reasonably necessary to prevent the harm. Further details are discussed below.

3.3.2 Corporate Liability under the Proviso to Article 26 of the Act Governing the Choice of Law in Civil Matters Involving Foreign Elements

If the proviso to Article 26 of the Act Governing the Choice of Law in Civil Matters Involving Foreign Elements applies, several potential choice-of-law rules may be invoked: (1) the law of the place where the damage occurred — in this case, applicable EU regulations, with conclusions as described above; (2) the law of the place where the victim purchased the product — in this case, U.S. law, with corporate liability as described above; and (3) the law of the victim's country of nationality — in this case, an EU Member State's law, under which the question of the manufacturer's liability will be assessed according to relevant EU norms, potentially resulting in significantly greater corporate responsibility and damages.

4 Corporate Liability Interlinked with Causation and the Burden of Proof

4.1.Causation and Strict Liability

After 1985, Directive 85/374/EEC enshrined the strict liability principle: victims need not prove the manufacturer's fault, but only a causal link between a product defect and the damage sustained; manufacturers are therefore liable for harm caused by defective products. This principle has been fully transposed into the domestic laws of all EU Member States. Consequently, absent a demonstrable lack of causal connection between the product and the consumer's health impairment, manufacturers will generally be held liable; they cannot avoid liability merely by invoking the absence of fault, although they may be exonerated if post-incident expert evidence establishes no causal relation to the product.

4.2. Legal Liability Arising from the Allocation of the Burden of Proof

4.2.1 Burden-of-Proof Rules in the European Union

The Product Liability Directive (Directive (EU) 2024/2853) [1], adopted in 2024, introduced a presumption-of-proof mechanism particularly tailored to situations of pronounced information asymmetry within the food supply chain[2]. The measure is intended to address victims' inability to prove highly technical causal links: under the traditional regime, plaintiffs must demonstrate the existence of a product defect, a burden that is especially onerous in transnational food chains because critical production processes, testing data, and internal quality-control records are typically controlled by corporations[3]. Consequently, the Directive permits provisional presumptions that a product is defective, or that a defect is causally linked to the harm, thereby increasing victims' prospects of prevailing in litigation.[4]

Accordingly, applying the Product Liability Directive (Directive (EU) 2024/2853) through choice-of-law rules would effectuate a reversal of the evidentiary burden, relieving consumers of the obligation to prove causation—a result that benefits consumers and disadvantages producers. Although the Directive primarily targets businesses that market products to, or whose products circulate within[5], EU Member States, its rationale is intended to inform domestic tort law across the Member States. Although the Directive applies to businesses that market products to EU Member States or whose products circulate within EU Member States, its spirit has become the basis for the allocation of the burden of proof in product liability cases.

The manufacturer mentioned in the preceding paragraph is the person who produces, manufactures, or processes the merchandise. Those, who attach the merchandise with the service mark, or other characters, signs to the extent enough to show it was produced, manufactured, or processed by them, shall be deemed to be the manufacturer.

If the production, manufacture, process, or design of the merchandise is inconsistent with the contents of its manual or advertisement, it is deemed to be defective.

The importer shall be as liable for the injury as the manufacturer."

[1] The revised Product Liability Directive (Directive on liability for defective products) was adopted by the European Parliament and the Council of the European Union on 23 October 2024 and published on 18 November 2024; it entered into force on the twentieth day following publication and is therefore currently in effect. See Taipei Economic and Trade Office in Surabaya, "EU Revision of the Product Liability Directive: Overview," 29 November 2024, p. 1, available at https://reurl.cc/9bQgQx. (as available on 18.5.2026)

[2] Article 7 of Directive (EU) 2024/2853.

[3] European Commission (2024), Guidance on the Application of Product Liability to Digital Products, Brussels, pp. 6-9.

[4] Quality Association, Beyond the Artificial Intelligence Act: Product Liability Directive and Artificial Intelligence Liability Directive, see https://reurl.cc/YknLMx (as available on 18.5.2026)

[5] Article 2 of Directive (EU) 2024/2853: "1. This Directive shall apply to products placed on the market or put into service after 9 December 2026."

4.2.2 Burden-of-Proof Rules in Taiwan's Jurisdiction

The presumption of causation and allocation of the burden of proof for product liability in our country are provided in Civil Code Article 191-1. If a manufacturer cannot prove that the production, manufacture, processing or design was free of defects, or that the damage was not caused by such a defect, or that it exercised due care to prevent the damage, then even if the manufacturer establishes that the addition of titanium dioxide was lawful under Taiwanese regulations, it will still be liable for compensation for health damage resulting from consumption.

4.2.3 Comparative Analysis of the Burden of Proof for Product Causation: EU versus Taiwan

From the foregoing food-safety provisions, it is evident that presumptions of causation and the burden of proof have been shifted onto manufacturers in both the EU and Taiwan; on this point the rules are the same and there is no divergence.

At present, both the EU and Taiwan regulate causation in product cases as a substantive matter by adopting a presumption-based liability regime. Thus, where harm arises from the normal use of a product—such as adverse health effects following consumption—the plaintiff need only demonstrate that the injury originated from the product, upon which the manufacturer is required to assume liability for damages. This,if the product was not used normally, the consumer bears responsibility.

If a manufacturer seeks to avoid liability for damages, it must adduce evidence that the adverse outcome resulted from climatic and environmental changes in Europe—such as food spoilage or the induction of unhealthy or allergic reactions following ingestion of additives—rather than from defects in the product's manufacture, formulation, or design. For example, titanium dioxide particles ingested in the European diet readily adsorb particulates in the gastrointestinal tract; climate change—by warming and acidifying the oceans (ocean acidification, OA)[1] and provoking alternating extremes of intense rainfall and drought—alters river runoff dynamics. During drought, heavy metals in soils and sediments may oxidize, and subsequent heavy rains can wash these mobilized metals into the sea in large quantities, thereby changing coastal salinity and pH.[2] When the oceans absorb excess carbon dioxide, seawater pH declines; in such more acidic conditions, heavy metals previously sequestered in marine sediments (e.g., cadmium [Cd], lead [Pb], mercury [Hg], arsenic [As]) undergo speciation changes, shifting from stable bound forms to more labile ionic forms, which are readily taken up by lower marine organisms such as algae and bivalves[3], and this leads marine organisms to accumulate greater loads of microplastics and heavy metals, producing the so-called "Trojan horse" effect[4]. This amplifies the bioaccumulation and biomagnification of toxins within the food web, increasing the risk that, when humans consume seafood, titanium dioxide particles enhance the uptake of marine pollutants and thereby raise the incidence of gastrointestinal inflammation or allergic reactions.[5] If such adverse outcomes stem from environmental contamination and consumer dietary practices rather than from spoilage or manufacturing defects in the food itself, the duty of care for safe consumption rests with the consumer. Given that the EU prohibits the use of titanium dioxide as a food additive, European consumers are expected to possess baseline knowledge and prudence; when purchasing or consuming foods legally containing titanium dioxide in other jurisdictions, they should avoid mixing such products in ways that could precipitate health risks. Consumption choices are therefore matters of individual autonomy and risk-bearing, not the responsibility of manufacturers.

4.2.4 Burden of Proof Where Physical Evidence Has Been Lost

If a victim requires physical evidence beyond proof of causation—for example, where the consumed product was discarded after the victim lost consciousness and thus cannot be produced—manufacturers and other actors in the supply chain have no inherent obligation to supply such items under ordinary rules. Under the 2024 Product Liability Directive, however, a court may compel the provision of relevant materials. To prevent plaintiffs from using such measures to obtain internal product information they otherwise could not access, plaintiffs should first establish prima facie evidence that the injury originated from the product—for instance, medical

---

[1] IPCC, IPCC Special Report on the Ocean and Cryosphere in a Changing Climate (SROCC), 2019, p450-453.

[2] Cooley, S., Schoeman, D., Bopp, L., Boyd, P., Donner, S., Ghebrehiwet, D. Y., Ito, S.-I., Kiessling, W., Martinetto, P., Ojea, E., Racault, M.-F., Rost, B., & Skern-Mauritzen, M., IPCC Sixth Assessment Repot--Climate Change 2022: Impacts, Adaptation and Vulnerability, Cambridge University Press, 2022, p.411-415, See https://www.ipcc.ch/report/ar6/wg2/ (as available 18.5.2026)

[3] Frank J. Millero, Ryan Woosley, Benjamin DiTrolio, & Jason Waters(2009), Effect of Ocean Acidification on The Speciation Of Metals In Seawater, Vol. 22, No. 4, December, See https://tos.org/oceanography/assets/docs/22-4_millero.pdf (as available 18.5.2026)

[4] Trevisan, R., et al.(2020) Pah Sorption To Nano plastics And The Trojan Horse Effect As Drivers Of Mitochondrial Toxicity And Pah Localization In Zebrafish, Environmental Science & Technology, vol. 54, No.15, p.9431-9439.

[5] For example, research has also reported that titanium dioxide nanoparticles exacerbate the bioaccumulation and toxicity of cadmium in mussels; see Canesi, L., et al., "Titanium dioxide nanoparticles aggravate the bioaccumulation and toxicity of cadmium in Mytilus galloprovincialis," Science of the Total Environment, 466, 2014, pp. 631–637.

records documenting treatment, clinical injury reports, purchase receipts, witness testimony, or widely reported instances of similar defects associated with the same product model on the market.

5 Conclusion (Management Recommendations For Business Owners)

In the EU, product liability arising from environmental change is governed by Article 7 of Rome II, under which the applicable law may be either the law of the country where the damage occurs or the law of the country in which the event giving rise to the damage occurs. In this case, the application of different connecting factors yields divergent outcomes: under the former the producer would be liable, whereas under the latter the producer would not be liable. Ordinary product-related torts are governed by Article 5 of Rome II, which applies either the law of the habitual residence (lex domicilei) or the rule of the manifestly closest connection.

When an action is brought before Taiwanese courts, Article 26 of the Act on the Application of Laws to Foreign-related Civil Relations governs choice of law: where the rule prescribes application of the law of the manufacturer's country, Taiwanese substantive civil law applies in the present case. However, if the proviso to Article 26 is applicable, the law of the place where the damage occurred or the law of the victim's nationality may instead lead to application of the relevant EU tort laws; and where the law of the place where the plaintiff purchased the product is the connecting factor, U.S. law may be applied. Thus, different applicable laws yield divergent outcomes as to corporate liability and damages.

Based on the foregoing analysis of this case, manufacturers and marketers must reassess their practices. With global travel and tourism facilitating the movement of consumers across jurisdictions, even where an ingredient is legally permitted in the jurisdiction of production or sale, R&D and production teams should proactively design formulations to meet the strictest international standards. This approach accords with the precautionary principle in risk governance: "Where there are plausible grounds for concern that a potentially dangerous effect on the environment or human health may be inconsistent with the high level of protection chosen, lack of full scientific certainty shall not be used as a reason for postponing cost-effective measures to prevent environmental degradation." Although compliance with stricter standards may raise production costs, those expenses are likely modest compared with potential liability awards (some of which can be enormous).

Concurrently, firms should implement Failure Mode and Effects Analysis (FMEA) during the food design stage, incorporating long-term cumulative toxicity risks into the FMEA assessment matrix. By applying Risk Priority Numbers (RPNs) to quantify subjective concerns, organizations can render latent risks visible and thereby inform managerial decision-making and remedial action.

Moreover, manufacturers should attend to principles of information symmetry and conspicuous notice. They should employ internationally recognized warning icons and prominently display in bold on product packaging the following: "This product is intended for sale and consumption only within [specified country/region]. Any risks arising from transporting and storing the product outside the authorized territory are the responsibility of the holder." Although such warnings will not entirely absolve legal liability, they constitute important evidence in litigation that the manufacturer fulfilled its duty to inform.

It is recommended that firms establish blockchain-grade traceability systems. When a product issue arises—such as the additive dispute described above—companies must be able, within minutes, to leverage their ERP (Enterprise Resource Planning) systems to trace a batch's raw-material suppliers, the specific production line that manufactured it, and the distributors to whom it was sold. Such capability permits containment of harm to discrete batches and enables precise attribution of responsibility in litigation.

Unforeseen events often occur beyond reasonable expectations. Accordingly, firms should transfer risk where appropriate by obtaining adequate insurance so that potential awards are shared with insurers. They should also establish a Product Launch Review Committee composed of R&D, Quality Assurance, Marketing, and Legal functions to identify and mitigate regulatory and corporate risks across jurisdictions, thereby enabling proactive prevention. In addition, companies ought to set aside dedicated legal and recall reserves to ensure that cash flow remains solvent when facing class actions or cross-border litigation.

References

Canesi, L., et al. (2014) Titanium dioxide nanoparticles aggravate the bioaccumulation and toxicity of cadmium in Mytilus galloprovincialis, Science of the Total Environment, 466, p.631-637.

Cooley, S., Schoeman, D., Bopp, L., Boyd, P., et al. (2022) IPCC Sixth Assessment Repot, "Climate Change 2022: Impacts, Adaptation and Vulnerability, Cambridge University Press", pp.411-415, See https://www.ipcc.ch/report/ar6/wg2/

European Commission(2016), Question number:EFSA-Q-2011-00348, Re-evaluation of titanium dioxide (E 171) as a food additive-- EFSA Panel on Food Additives and Nutrient Sources added to Food (ANS), First published:14 September 2016, EFSA Journal, Vol.14, No. 9 , p.68, See https://efsa.onlinelibrary.wiley.com/doi/10.2903/j.efsa.2016.4545

European Commission (2024), Guidance on the Application of Product Liability to Digital Products, Brussels, pp. 6-9.
FAF Panel EFSA Panel on Food Additives and Flavourings (FAF) (2021), Scientific Opinion on the safety assessment of titanium dioxide (E171) as a food additive, EFSA Journal, May, pp.1,9,73-87, https://doi.org/10.2903/j.efsa.2021.6585
Frank J. Millero, Ryan Woosley, Benjamin DiTrolio, & Jason Waters(2009), Effect of Ocean Acidification on The Speciation Of Metals In Seawater, Vol. 22, No. 4, December, See https://tos.org/oceanography/assets/docs/22-4_millero.pdf
IPCC, IPCC Special Report on the Ocean and Cryosphere in a Changing Climate (SROCC), 2019, p450-453.
Niqhat Fatma Rizvi, Anoop Kumar, Rashmi Chandra, and Shoebul Haque (2024), Synthetic Food Colourants: Applications, Regulatory Perspectives, and Potential Health Implications, World Journal of Biology Pharmacy and Health Sciences (WJBPHS), Vol.17, No. 1, January, p.63. See DOI: 10.30574/wjbphs.2024.17.1.0068
Regulation (EU) No 1215/2012 of the European Parliament and of the Council of 12 December 2012 on jurisdiction and the recognition and enforcement of judgments in civil and commercial matters (Brussels I Recast), OJ L 351/1, 2012.
Safety assessment of titanium dioxide (E171) as a food additive, EFSA Journal, Vol. 19, No.5, p.78, https://efsa.onlinelibrary.wiley.com/doi/epdf/10.2903/j.efsa.2021.6585
Standards for the Scope of Use, Maximum Limits, and Specifications of Food Additives (amended on 7 May 2025 [R.O.C. Year 114]), Annex 1: Scope of Use and Maximum Levels of Food Additives – Category IX: Colorants. Available at: chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://db.lawbank.com.tw/GetFile.ashx?pfid=0000297834
Taipei Economic and Trade Office in Surabaya, "EU Revision of the Product Liability Directive: Overview," 29 November 2024, p. 1, available at https://reurl.cc/9bQgQx.
Taiwan Food and Drug Administration (TFDA), Ministry of Health and Welfare, preannounced a draft revision of the "Analytical Methods and Specifications for Titanium Dioxide as a Food Additive" on 23 March 2026. https://www.fda.gov.tw/tc/newsContent.aspx?cid=5072&id=31461
Trevisan, R., et al.(2020) Pah Sorption To Nano plastics And The Trojan Horse Effect As Drivers Of Mitochondrial Toxicity And Pah Localization In Zebrafish, Environmental Science & Technology, vol. 54, No.15, p.9431-9439.
Quality Association, Beyond the Artificial Intelligence Act: Product Liability Directive and Artificial Intelligence Liability Directive, see https://reurl.cc/YknLMx

IFKAD 2026

Organized by:
Corvinus University of Budapest
LUM University Giuseppe Degennaro
Arts for Business

Associated Partners:
Emerald Publishing
International Association for Knowledge Mngt
Institute of Knowledge Asset Mngt
AiIG - Assoc. Italiana di Ingegneria Gestionale
National Chengchi University
Industry+Innovation LABoratories

Book of Papers IFKAD 2026
ISBN: 978-88-96687-20-8